12

THE ASSESSMENT OF A JOINT PROBABILITY DISTRIBUTION ON AN EVOLVING DYNAMIC UNCERTAIN PROCESS

STANFORD UNIVERSITY

Dennis M. Buede

ADVANCED ARPA DECISION TECHNOLOGY PROGRAM

CYBERNETICS TECHNOLOGY OFFICE
DEFENSE ADVANCED RESEARCH PROJECTS AGENCY
Office of Naval Research • Engineering Psychology Programs

The objective of the Advanced Decision Technology Program is to develop and transfer to users in the Department of Defense advanced management technologies for decision making. These technologies are based upon research in the areas of decision analysis, the behavioral sciences and interactive computer graphics. The program is sponsored by the Cybernetics Technology Office of the Defense Advanced Research Projects Agency and technical progress is monitored by the Office of Naval Research — Engineering Psychology Programs. Participants in the program are:

Decisions and Designs, Incorporated
Harvard University
Perceptronics, Incorporated
Stanford Research Institute
Stanford University
The University of Southern California

Inquiries and comments with regard to this report should be addressed to:

Dr. Martin A. Tolcott
Director, Engineering Psychology Programs
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

or

LT COL Roy M. Gulick, USMC
Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209

RESEARCH REPORT No. EES DA-77-5

(6) # THE ASSESSMENT OF A JOINT PROBABILITY DISTRIBUTION ON AN EVOLVING DYNAMIC UNCERTAIN PROCESS.

(9) Research rept.

by elvin

(10) Dennis M. Buede

DECISION ANALYSIS PROGRAM

Professor Ronald A. Howard
Principal Investigator

Sponsored by

Advanced Research Projects Agency
ARPA Order No. 3052
Under Sub-contract from
Decisions and Designs, Incorporated

and

National Science Foundation, NSF Grant ENG 76-81056

(12) 150 p.

(11) Sepetember 1977

(15) N00014-76-C-0074, ARPA Order-3052

DEPARTMENT OF ENGINEERING-ECONOMIC SYSTEMS
Stanford University
Stanford, California 94305

ABSTRACT

An evolving dynamic uncertain process is one of several complicating factors in the analysis of decisions for which time is a critical element. A process is uncertain when its output cannot be specified deterministically; dynamic when it and its output change with the passage of time; and evolving when the output is revealed as time progresses. The decision maker's uncertainty on the dynamic random variable that characterizes such a process must be adequately represented in any decision analysis. The direct assessment of this uncertainty is preferable to further modeling when the expert's knowledge of the variable is roughly equivalent to his knowledge of its components.

This dissertation addresses the problem of performing the assessment of a joint probability distribution on a dynamic random variable in a practical manner. The first approach to the problem is to examine multivariate named distributions; they are found to be too rigid in their structure to be good approximations of the decision maker's joint probability distribution, in general.

Since multivariate named distributions are too restrictive, the second approach is to develop a more flexible method of assessing a joint probability distribution on a dynamic

random variable. This method is a general model of the decision maker's use of new information to update his probabilistic beliefs. A mathematical framework for specifying a posterior distribution in terms of the prior distribution and the revealed information is introduced here. This framework provides the analyst with a great deal of flexibility for approximating any joint distribution the decision-maker might have.

TABLE OF CONTENTS

ABSTRACT	ii
LIST OF FIGURES	vi
LIST OF TABLES	viii
ACKNOWLEDGMENTS	ix
CHAPTER 1. INTRODUCTION AND SUMMARY OF RESULTS	1
1.1 Introduction	1
1.2 The Problem	3
1.3 Summary of Results	10
1.4 Related Work	12
CHAPTER 2. THE METHOD OF MULTIVARIATE NAMED DISTRIBUTIONS	19
2.1 Introduction and Summary	19
2.2 The Multivariate Normal Distribution	20
2.3 The Multivariate Lognormal Distribution	24
2.4 The Multivariate Student Distribution	27
CHAPTER 3. THE METHOD OF UPDATING THE PRIOR DISTRIBUTIONS	32
3.1 Introduction	32
3.2 The Method	34
3.3 Relaxing the Assumption of Normal Distributions	50
3.4 Sources of Error	57
CHAPTER 4. THE ASSESSMENT PROCEDURE FOR THE METHOD OF UPDATING THE PRIOR DISTRIBUTIONS	61
4.1 Introduction	61
4.2 The Preliminary Assessment Phase	63
4.3 The Preliminary Analysis Phase	66
4.4 The Final Assessment Phase	78
4.5 The Final Analysis Phase	80
4.6 Verification by the Decision maker	82
CHAPTER 5. THE ASSESSMENTS	86
5.1 Introduction	86
5.2 An Overview of the Assessments	86
5.3 The Assessment of the Price of U_3O_8	92

CHAPTER 6. CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH 111

6.1 Conclusions 111
6.2 Suggestions for Further Research 112

APPENDIX A. NOTATION 114

APPENDIX B. PROOF OF EQUATION 2.7 118

APPENDIX C. THE POSTERIOR MULTIVARIATE STUDENT DENSITY FUNCTION 120

APPENDIX D. A DESCRIPTION OF A GENERAL METHOD FOR UPDATING 123

APPENDIX E. PROOF OF EQUATIONS 3.21 AND 3.23 132

REFERENCES 133

DISTRIBUTION LIST 137

DD FORM 1473 140

LIST OF FIGURES

Figure

Figure		Page
1.1	The consumption/investment decision	4
2.1	The effect of the degree of freedom parameter n on the posterior variance of $x(2)$ given $x(1)$	30
3.1	The method of updating the prior distributions	33
3.2	A graphical representation of the median updating algorithm	39
3.3	A first approximation - translation	44
3.4	A graphical representation of $g_{0.9}$	45
3.5	A second approximation - rotation	49
3.6	Possible shapes of assessed probability distributions	51
4.1	The assessment procedure	62
4.2	An example of a graph of $x(i\|i-1)_{0.5}$ versus $x(i-1)$ and $x(i-2)$	65
4.3	The preliminary assessment phase	67
4.4	Possible types of interaction between $x(i-1)$ and $x(i-2)$ on the posterior median	72
4.5	The effect of the nonlinear multiplicative factor	74
4.6	The preliminary analysis phase	79
4.7	The final assessment phase	81
4.8	The final analysis phase	83
4.9	The verification by the decision maker	85
5.1	The prior fractiles on the price of U_3O_8	95
5.2	The expert's preliminary graphs of the posterior median versus $x(i-1)$ and $x(i-2)$	96

5.3 $g_{0.5}$ versus $f(1)$ for $i=2$ 102

5.4 $g_{0.5}$ versus $f(2)$ for $i=3$ 104

5.5 $g^*_{0.9}$ versus $f(1)$ for $i=2$ 106

5.6 $g^*_{0.9}$ versus $f(2)$ for $i=3$ 107

D.1 A graphical representation of g 124

D.2 Some properties of g_f 127

D.3 A graph of g versus f for the bivariate normal, $f(1) = 0.5$ 128

D.4 A graph of g versus f for the bivariate normal, $f(1) = 0.25$ 130

LIST OF TABLES

Table

5.1 A Summary of the Assessments 88

5.2 The Assessment Errors 90

5.3 The Approximate Prior Distributions 97

5.4 The Given Information for the Assessed Posterior Distributions 99

5.5 8 Assessed Posterior Distributions 100

5.6 Comparison of Assessed and Approximate Posterior Distributions 109

ACKNOWLEDGMENTS

My deepest appreciation goes to Professor Ronald A. Howard, whose advice and support were crucial to my education and research efforts. I am also indebted to Dr. James E. Matheson, Dr. Peter A. Morris, and Professor William K. Linvill for reviewing this manuscript and providing many useful suggestions in numerous discussions during the final stages of my research.

There are numerous fellow students who contributed a great deal to my overall education. I would especially like to thank those individuals who gave their time and thoughts during the assessments that were critical to this research: Gary Frisvold, Thomas Keelin, Daniel Owen, Steve Powell, and Richard Urso. I also wish to thank Tom Banet, Allen Grum and Bruce Judd for their contributions to my initial efforts in the world of academic research.

My wife, Anne, provided much more support during this period than was warranted and for this I am very grateful.

Decisions and Designs, Incorporated graciously provided support for the final typing of this manuscript. My thanks go to Kathy McKnew and Marilyn Pickett for their diligent typing services.

This research was supported by the Advanced Research Projects Agency of the Department of Defense (monitored by the Office of Naval Research) under Contract N00014-76-C-0074 and under Decisions and Designs, Inc., Subcontract 75-030-0713, and by the National Science Foundation, under Grant NSF ENG-76-81056.

CHAPTER 1

INTRODUCTION AND SUMMARY OF RESULTS

1.1 Introduction

In many decision problems one or more of the uncertain variables is dynamic. By dynamic we mean the variable changes value with the passage of time. These changes in value are often critical in determining the best alternative for the decision maker. Decision analysis is the logical approach that is most helpful to the decision maker for finding the best decision. (See [14] for a discussion of decision analysis.) However, the assessment of the joint probability distribution needed to capture the decision maker's uncertainty on a dynamic random variable (drv) is very difficult. This difficult assessment is directly related to the characteristics of the process that is generating the values of the drv. This process is evolving, dynamic, and uncertain; uncertain because it cannot be specified deterministically by the decision maker; dynamic because it and its output, the drv, change with the passage of time; and evolving because values of the drv are disclosed as time passes. The assessment of the joint probability distribution on the values of a drv in the

time periods of interest to the decision maker is the topic of this research.

The definition of probability adopted in this research will be the Bayesian definition (known as subjectivistic or personalistic). This view asserts that probability depends on the state of information of the individual. Thus it is conceivable that people with different states of information will have different probabilities for the same set of events. Discussions of this and other definitions of probability can be found in references [9, 20, 27, 30]. Inferential notation will be used to denote probabilities. This notation is discussed at the end of Appendix A.

1.2 The Problem

An example might be useful to demonstrate the difficulty inherent in assessing the joint probability distribution on a drv. Consumption/investment decisions are faced by nearly every individual or household. To find the best decision strategy, these decisions must be considered sequentially since each decision affects all of the others. For these problems it is helpful to consider N discrete time periods, at the beginning of which some income is received and during which consumption takes place. The consumption/investment decision $c(i)$ is made at the beginning of each period, just after the value of that period's income $x(i)$ is revealed. For this example we will assume that the difference between the decision maker's wealth in period i and his consumption in period i is invested in a safe asset at a fixed rate. The crucial variable in this decision is the decision maker's income in each time period since each of these values affects all of the consumption decisions. Unfortunately each value is uncertain. But even worse, there is a significant amount of (probabilistic) dependence between the decision maker's income in one time period and his income in the other time periods. This sequential decision is illustrated in Figure 1.1 with a decision tree.

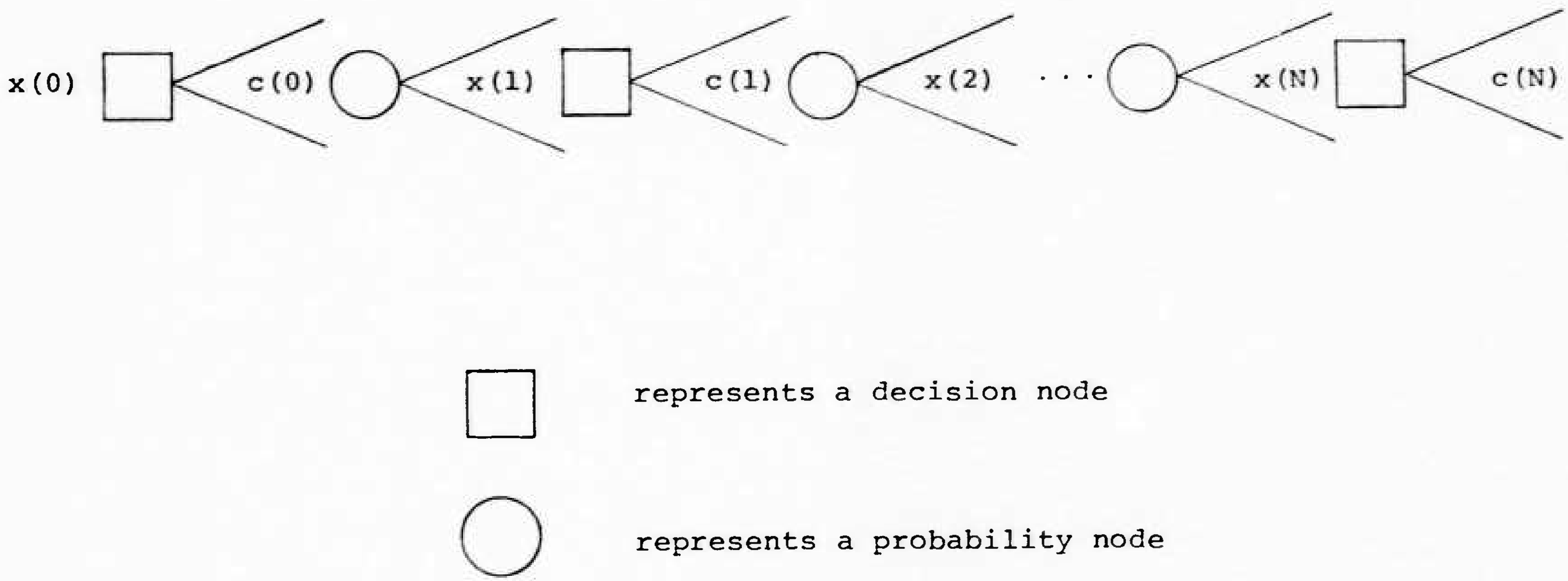


Figure 1.1 The consumption/investment decision.

In order to calculate the optimum consumption for the present period, a dynamic programming problem must be solved. (For a review of dynamic programming see [3].) This solution necessitates the calculation of the optimum consumption in each time period for all possible values of income and consumption in the previous time periods. Thus a posterior probability distribution for each period's income is needed. These probability distributions must be conditioned on every possible previous path of the income variable. These probability distributions

$$\{x(1) \mid \xi\}$$

$$\{x(2) \mid x(1), \xi\}$$

$$\cdots$$

$$\{x(N) \mid x(N-1), x(N-2), \ldots, x(1), \xi\}$$

are sufficient to specify the joint probability distribution on $x(1), x(2), \ldots, x(N)$. It is clear that direct assessment of these probability distributions would be excessively difficult, if not impossible, even if the random variable were discretized. For example, if we discretized income into three regions, we would have to assess $\sum_{j=0}^{N-1} 3^j$ distributions assuming no conditional independencies. Even if there were some conditional independencies, this would still be a large number for N greater than five or six.

Assessment problems of this type plague the solution of many dynamic decision problems, even some that are not sequential decisions like the previous example. The probabilistically dependent structure of a drv is the rule rather than the exception for these decisions. That is, the value of $x(j)$ will provide some information about the value of $x(i)$, $(j < i)$.

Now it is important that we specify more carefully the type of drv being dealt with in this research. First, we must state that the results of this research concerning the assessment of a joint probability distribution on a drv are to be applied during the probabilistic phase [14, p. 284] of a decision analysis. That is, we have completed the deterministic phase of the analysis and have a deterministic model of the decision that identifies the crucial uncertain variables whose probability distributions are now to be assessed. Thus we have decided that we should do no further modeling of the many aspects of the decision, only assessment. Our research here is concerned with how the assessment of a joint probability distribution on a drv can be done in a practical yet rigorous manner.

Besides the probabilistic dependence of the periodic values of the drv there are three other characteristics of a drv which will complete its specification for this research. The first is the independence between the drv and all decision variables. If this were not satisfied, the assessments would also have to be made for all values of the decision variables. Decision analysts have found that most practical problems can be structured in such a way that there is probabilistic independence between decision and state variables. Thus this condition is usually satisfied in practice.

The second characteristic of a drv is that it is probabilistically independent of all of the other variables in the decision. In other words, the only information that will change the decision maker's probability distribution on the value of a drv in a given time period is the values of the drv in the previous time periods. Decision analysts have also found that many problems can be structured so that this condition is met. It may not be possible to achieve in all decision problems, but we will have a better idea of how to deal with this situation when we have found a convenient way to assess the joint distribution when it is satisfied.

The inability of the decision maker to assume that the evolving dynamic uncertain process (which is generating the values of the drv) is stationary, is the third characteristic. (For a discussion of stationary processes see [4, 19].) If this process were stationary, there would be a set of deterministic relationships among the periodic values of the drv such that a small number of probability distributions would describe all of the uncertainty on the drv. Thus the state-of-the-art can currently handle the assessment of a joint distribution on a drv being generated by a stationary process. However, in this research the generation process of the drv is assumed to be changing over time in a stochastic manner and is therefore nonstationary. An important aspect of a nonstationary process is that the more recent an observation, the more influential it will be in updating the probability distributions on future values of the drv. This method can also be used when the drv is being generated by a stationary process. However, nonstationary processes are the ones that we are least equipped to handle.

Now that the problem has been defined, a few words about the important aspects of its solution should be presented. The principal aspects of any probability assessment technique are its theoretical foundations,

its informational requirements, and its flexibility for adequately approximating any probabilistic beliefs. First, the solution should be theoretically based on the concepts of probability theory. The probabilistic updating process on the value of $x(i)$ should be well defined in terms of the revealed information, $x(1), x(2), \ldots, x(i-1)$, being used to transform the prior, $\{x(i) | \xi\}$, to the posterior, $\{ x(i) | x(i-1), \ldots, x(1), \xi \}$.

The informational requirements of the decision maker should be kept at a practical level. Obviously the less information he has to supply in order for the analyst to specify his joint probability distribution, the better the method. Typically, the analyst will have to make a trade-off between obtaining more information, which increases the cost of the analysis, and settling for less precision by making do with what he currently has.

Finally the most discriminating attribute of the assessment methods investigated was their flexibility for approximating any of the many possible probabilistic beliefs that a decision maker might have. Since the mean and variance of a probability distribution are often

its most important descriptors, the relationships between the posterior mean (or median) and variance of $x(i)$ and the revealed values of $x(1), x(2), \ldots, x(i-1)$ were taken to be indications of the flexibility of the approximation methods. That is, the number of different types of functional relationships between the posterior mean (median) and variance and the revealed path that a method could exhibit is a measure of its flexibility. This will be discussed in much greater detail in later chapters.

1.3 Summary of Results

The major contribution of this research is a practical method for approximating a decision maker's joint probability distribution on a dynamic random variable. This method includes both a mathematical framework for specifying a joint distribution and an assessment procedure. The mathematical framework is original to this research and has proven in seven assessments to be robust. The assessment procedure uses state-of-the-art encoding techniques to gather the necessary information needed by the analyst to dress the mathematical framework.

Another contribution is a formalization of the difficulties inherent in a solution to this problem and

a general discussion of several possible solutions.

Chapter 2 provides a summary of the three most useful multivariate named distributions--the normal, lognormal, and Student--that can be used to characterize a decision maker's uncertainty on a drv. The contribution here is an evaluation of these distributions for use in this particular way.

The mathematical framework of the method for approximating a joint distribution on a drv is completely presented in chapter 3. Section 1 is a brief introduction to the general aspects of the method. The nature of this framework is completely developed in the second section, including the assumptions and mathematical details. The third section then discusses a relaxation of one of the assumptions in order to make the method more generally applicable. Section 4 identifies and discusses the sources of error in this approximation technique.

The assessment procedure is presented, step by step, in chapter 4. This procedure uses recognized encoding techniques and mathematical manipulations to ascertain systematically the form and parameters of the proper mathematical framework to approximate the decision maker's uncertainty.

Chapter 5 then presents a discussion of seven assessments that we performed. A general overview is provided in the second section to highlight their characteristics and associated errors. One of our assessments is then presented in detail to illustrate the workings and usefulness of this assessment technique.

Chapter 6 presents our conclusions and some thoughts on areas for future research.

1.4 Related Work

There has been no thorough treatment of this problem previous to this research. In the past, assessments of these type have been simplified by assumptions, such as linearity or stationarity.

Bather [2] and North [24] have investigated a similar problem. They assumed that the decision is not directly affected by the values of the drv $\underline{x}$ but by a parameter $\underline{\theta}$ which describes the generation of the drv. However, the decision maker is only able to observe values of $\underline{x}$, the drv. The generation process of $\underline{\theta}$ is assumed to be Markov in the sense that the distribution $\{\theta(i)|\theta(i-1),\xi\}$ is probabilistically independent of the previous values of $(\theta(1),\theta(2),...,\theta(i-2))$ for i between 2 and N. This Markov assumption is impractical for the general drv and was

not considered for this research. Also, the actual value of the drv will be observed as time evolves, so this problem does not seem to be as directly relevant as the one being investigated in this research.

Several other possible solutions have been investigated by the author. One possible solution involves the assessment of the probabilities on the paths that the drv will take over time. It will be easier to describe this method if we assume time is continuous, and we are interested in the values of the drv in the interval of time $(0,\tau)$. There are an infinite number of paths the drv could follow in this interval of time, and each path can be described by a function $x(t)$ which need not be continuous or restricted in any other way. Of course we could assume some general restriction, but it will become clear that this will not really solve the problems inherent in the method. (Note that the paths cannot be restricted to be finite in number because then there will always be some fraction of the path which could be revealed such that the future is certain. This is not a sufficiently good representation of reality to be useful.) The approach would then be to encode a probability distribution on the possible paths x may take over time. Then after a certain amount of time has passed, say $\tau'<\tau$, and a partial path has been revealed, the analyst can use this probability distribution and the probability calculus to find the posterior distribution on the paths of

x from time τ' to τ given the partial path from 0 to τ'. If some sort of conjugate distribution exists, this operation could be done analytically. The assessed distribution on the paths of x provides all of the probabilistic information needed to determine the probability distribution on x at a specific time t given any state of information.

Unfortunately, we are not equipped either mathematically or practically to handle information in the form of paths. To begin with, there are an infinite number of paths that the drv can follow between any two points in time. However, this is not the half of it (figuratively speaking). The number of integers and the number of real numbers between 0 and ∞ represent two different orders of infinity. (See [11] for a complete discussion of this.) However, the number of functions that map t into x over a finite time interval is an even higher order of infinity than these two. (For a proof of this see [23].) A passage from Georgescu-Roegen [12, pp. 76-77] sums up this situation:

> "As we say in mathematics, the continuum of the real number system forms only a simple* infinity.

*Underlined text denotes italics in original.

> The suggestion, natural at this juncture, of using more than one real number, i.e., a vector, for labeling qualities would still not reduce quality to number. For, as set theory teaches us, no matter how many coordinates we add, no set of vectors can transcend simple infinity in power. ... the next higher cardinal number mathematics has been able to construct after that of the arithmetical continuum is represented by the set of functions of a real variable, i.e., by a set of forms. Clearly then, forms cannot be numbered."

Questions one might be tempted to ask at this point are: Is it really this mathematical complexity of infinity that renders this method impractical? After all, we have been encoding probability density functions on continuous random variables without any trouble. Is not the problem one of giving a cardinal or even ordinal labeling to these paths on functions of x? If this could be done, the method would be usable. For then the paths could be labeled in an ordinal way, and comparative questions could be asked about these paths. The point being made by Georgescu-Roegen is that the infinity of paths cannot be numbered or labeled by any finite set of continuous numbers. Thus, until mathematicians have devised a labeling procedure for sets whose numbers of elements belong to his higher order of infinity, the method is impractical.

Even if such a labeling procedure were available, it would not be useful unless people were sufficiently familiar with it to answer the type of questions asked in encoding sessions. For instance, we can assume discrete time periods, and the problem of the higher order of infinity vanishes. Now we can use a vector to label the past values of the drv, namely the vector of its previous values. However, we still have the problem of being able to label systematically the infinite number of revealed vectors in terms such that people can answer questions like: "Is this set of paths as likely as that set"? or "Would you rather bet on the blue area of the wheel of fortune or the likelihood that one of the paths in ... set of paths will actually occur"? Thus, even if we assume discrete time and can avoid the mathematical problems of higher orders of infinity, this method does not appear to be of practical value because people do not think in terms of sets of paths and have a difficult time answering these types of questions. Even assumptions about the admissable shape of the paths, such as continuity, do not dissolve this very practical problem of human familiarity.

Another general area of research that could be applied as a solution to this problem is that of linear system theory. (See [7, 18].) This theory consists of linear models of dynamic systems with various degrees of uncertainty ranging from certainty to uncertainty of the parameters as

well as the exact form of the model. The most used part of this theory with uncertainty consists of a known linear model with known parameters and known additive noise. However, the assumptions made for this model imply the equivalent of a multivariate normal distribution on the drv. This is one of the methods which was researched and which will be described in Chapter 2. There are characteristics about the multivariate normal distribution which make it useful to control engineers but unfortunately not very useful for the topic of this research. This is the type of uncertainty that is often used in dynamic programming problems.

Another subset of linear system theory postulates a linear model with uncertain parameters and uncertain additive white noise. This problem has been examined in several places [21, 25, 31], and was thoroughly investigated as a solution to this problem by the author [5]. However, there are numerous mathematical and practical problems which make this approach infeasible. First, this approach presupposes that the decision maker can provide the analyst with a prior joint probability distribution on the parameters of this model. These parameters do not exist and have no intuitive meaning to the decision maker. Few assessments of this type have been attempted and none have been reported successful.

Second, the flexibility of this method is quite restricted. The relationships between the posterior median and standard deviation and the revealed path are very limited.

The development of sufficient statistics for specifying the joint distribution was also investigated [6]. Rather than defining the posterior distribution $\{x(i)|x(i-1),\ldots,x(1),\xi\}$ in terms of the revealed path $x(1)_\phi,\ldots,x(i-1)_\phi$, we attempted to develop a set of sufficient statistics that adequately summarized this information. However, this approach also failed because it did not have the flexibility to capture many different types of probabilistic beliefs.

Finally, an axiomatic approach was investigated. This approach attempted to map one or more sets of axioms that the decision maker might believe about any stochastic process into the subspace of possible joint probability distributions on the process. The complexity of the joint probability distributions encountered in practice was partially responsible for keeping this approach on the blackboard. An additional complication was the number of axiomatic structures needed to provide a mutually exclusive and nearly collectively exhaustive covering of the set of possible beliefs. Some examples of the possible beliefs about the uncertainty on a drv are discussed in chapter 4, section 3.

CHAPTER 2

THE METHOD OF MULTIVARIATE NAMED DISTRIBUTIONS

2.1 Introduction and Summary

In this chapter we will thoroughly examine and evaluate the most obvious method for approximating the decision maker's joint probability distribution on a drv that characterizes an evolving dynamic uncertain process. This is probably the most direct solution to this problem in terms of information required and ease of assessment. However, since these named distributions are so structured, we would expect their flexibility to be quite restricted. By flexibility we mean the ability of one method to approximate adequately any of the many possible probabilistic beliefs that a decision maker might have.

The multivariate named distributions examined during this research were the following: the normal, lognormal, Student, inverted Dirichlet (beta), Pareto, and several gamma's. The major source of information on these distributions is [16]. The multivariate normal and Student distributions are also discussed in [10,25] while the multivariate lognormal is presented in [17]. All but the normal, lognormal, and Student distributions are characterized by complicated joint, prior, and

posterior density and cumulative functions. Therefore only these three multivariate distributions were considered useful and are discussed in this chapter.

Sections 2, 3, and 4 of this chapter discuss the characteristics of the normal, lognormal and Student distributions, respectively.

2.2 The Multivariate Normal Distribution

The N-dimensional multivariate normal distribution is characterized by an N-dimensional vector of means $\underline{\mu}$ and a NxN-dimensional covariance matrix $\underline{\Sigma}$:

$$\underline{\mu} = \begin{bmatrix} \mu(1) \\ \mu(2) \\ \vdots \\ \mu(N) \end{bmatrix} \tag{2.1(a)}$$

$$\underline{\Sigma} = \begin{bmatrix} \sigma(1)^2 & \rho_{12}\sigma(1)\sigma(2) & \cdots & \rho_{1N}\sigma(1)\sigma(N) \\ \rho_{12}\sigma(1)\sigma(2) & \sigma(2)^2 & \cdots & \rho_{2N}\sigma(2)\sigma(N) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1N}\sigma(1)\sigma(N) & \rho_{2N}\sigma(2)\sigma(N) & \cdots & \sigma(N)^2 \end{bmatrix} \tag{2.1(b)}$$

The covariance matrix is symmetric and positive-definite. Thus, there are only $\frac{1}{2}(N^2+N)$ distinct elements of the

matrix. The N-dimensional joint density function of the multivariate normal is

$$\{\underline{x}|\xi\} = f_n^{(N)}(\underline{x}|\underline{\mu},\underline{\Sigma})$$

$$= (2\pi)^{-\frac{N}{2}}|\Sigma|^{-\frac{1}{2}}\exp(-\tfrac{1}{2}(\underline{x}-\underline{\mu})'\underline{\Sigma}^{-1}(\underline{x}-\underline{\mu})). \tag{2.2}$$

To determine the prior and posterior distributions of the multivariate normal we must first partition the vectors and matrices as follows:

$$\underline{x} = (x(1),\ldots,x(q)|x(q+1),\ldots,x(N))' = (\underline{x}_1',\underline{x}_2')',$$

$$\underline{\mu} = (\mu(1),\ldots,\mu(q)|\mu(q+1),\ldots,\mu(N))' = (\underline{\mu}_1',\underline{\mu}_2')',$$

$$\underline{\Sigma} = \begin{bmatrix} \underline{\Sigma}_{11} & \underline{\Sigma}_{12} \\ \underline{\Sigma}_{21} & \underline{\Sigma}_{22} \end{bmatrix} \quad \text{where } \underline{\Sigma}_{11} \text{ is qxq,} \tag{2.3}$$

$$\underline{P} = \underline{\Sigma}^{-1} = \begin{bmatrix} \underline{P}_{11} & \underline{P}_{12} \\ \underline{P}_{21} & \underline{P}_{22} \end{bmatrix} \quad \text{where } \underline{P}_{11} \text{ is qxq.}$$

Note that since $\underline{\Sigma}$ is the covariance matrix, $\underline{P}$ can be interpreted as the precision matrix. Then the general prior and posterior distributions can be written as:

$$\{\underline{x}_2|\xi\} = f_n^{(N-q)}(\underline{x}_2|\underline{\mu}_2,\underline{\Sigma}_{22}) \tag{2.4}$$

$$\{\underline{x}_2|\underline{x}_1,\xi\} = f_n^{(N-q)}(\underline{x}_2|\underline{\mu}_2+\underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}(\underline{x}_1-\underline{\mu}_1),(\underline{\Sigma}_{22}-\underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}\underline{\Sigma}_{12}))$$

$$= f_n^{(N-q)}(\underline{x}_2|\underline{\mu}_2-\underline{P}_{22}^{-1}\underline{P}_{21}(\underline{x}_1-\underline{\mu}_1),P_{22}^{-1}). \quad (2.5)$$

See [25] for a proof of this. Note that all prior and posterior distributions of a multivariate normal are either multivariate or univariate normal distributions.

It is important to note at this point that an implicit assumption of this multivariate distribution is the following: the information that is contained in the revealed path (the values of $x(1),x(2),...,x(i-1)$) and that is to be used to update the prior distribution on $x(i)$ can be totally described by the covariances between the periodic values taken two at a time. In some cases this can be too great a simplification of reality. For these cases this method would not yield a very satisfactory approximation of the decision-maker's joint probability solution. None of the assessments we did would have been adequately approximated by this distribution.

In order to determine the flexibility of the multivariate normal distribution we must examine the relationships between the posterior mean and variance of $x(i)$ and the revealed path, $x(1),x(2),...,x(i-1)$. The posterior mean

$$\langle\underline{x}_2|\underline{x}_1,\xi\rangle = \underline{\mu}_2+\underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}(\underline{x}_1-\underline{\mu}_1) \quad (2.6)$$

is a linear combination of the revealed values of the drv. This is the simplest nontrivial approximation one could assume. In fact it suggests the assessment procedure for the multivariate normal distribution described in [15]. It should be noted though that there is a restriction placed on this linear relationship between the posterior mean and a revealed value of the drv. This is that the posterior mean of $x(i)$ given that $x(j)$ $(j<i)$ equals its prior mean is equal to the prior mean of $x(i)$:

$$\langle x(i) | x(j) = \langle x(j) | \xi\rangle, \xi\rangle = \langle x(i) | \xi\rangle. \qquad (2.7)$$

In fact it is proven in Appendix B that this relationship is true whenever the posterior mean is a linear combination of the variables, regardless of the specific joint probability distribution.

The posterior variance

$$\overset{v}{<}\underline{x}_2 | \underline{x}_1, \xi\rangle = \underline{\Sigma}_{22} - \underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}\underline{\Sigma}_{12} \qquad (2.8)$$

is a constant and therefore independent of the revealed path. This was a particularly poor approximation for all of the assessments we did. There is an even greater limitation here though. It is that once the parameters of the multivariate normal distribution have been determined by the above-mentioned assessment procedure, the posterior variances are all determined. In other words,

once the decision maker has indicated how his posterior means should be updated, he has no freedom for specifying how the other characteristics of his posterior distributions should be updated. In fact, he may find that for some seemingly reasonable correlation coefficients, the covariance matrix is not positive-definite.

These characteristics of the multivariate normal distribution indicate that it has very little flexibility for approximating a decision maker's uncertainty on a drv and could only be used as a crude first approximation. However, it is the approach used quite often in the dynamic programming literature.

2.3 The Multivariate Lognormal Distribution

The N-dimensional multivariate lognormal distribution on $\underline{z}$ is obtained via the transformation:

$$x(i) = \ln(z(i)), \quad (i=1,2,\ldots,N) \tag{2.9}$$

on $\underline{x}$ when the joint distribution on $\underline{x}$ is multivariate normal. The multivariate lognormal density function is

$$\{\underline{z}|\xi\} = f_{ln}^{(N)}(\underline{z}|\underline{\mu},\underline{\Sigma})$$

$$= (2\pi)^{-\frac{N}{2}}|\underline{\Sigma}|^{-\frac{1}{2}}(\prod_{i=1}^{N} z(i))\cdot \tag{2.10}$$

$$\exp(-\tfrac{1}{2}(\ln(\underline{z})-\underline{\mu})'\,\underline{\Sigma}^{-1}(\ln(\underline{z})-\underline{\mu})),$$

$$z(i)>0,\ (i=1,2,\ldots,N).$$

Using the partitions of section 2, the prior and posterior distributions on $\underline{z}_2$ can be derived (see [17]):

$$\{\underline{z}_2|\xi\} = f_{ln}^{(N-q)}(\underline{z}_2|\underline{\mu}_2,\underline{\Sigma}_{22}) \tag{2.11}$$

$$\{\underline{z}_2|\underline{z}_1,\xi\} = f_{ln}^{(N-q)}(\underline{z}_2|\underline{\mu}_2+\underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}(\ln(\underline{z}_1)-\underline{\mu}_1),$$

$$(\underline{\Sigma}_{22}-\underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}\underline{\Sigma}_{12})) \tag{2.12}$$

$$= f_{ln}^{(N-q)}(\underline{z}_2|\underline{\mu}_{2|1},\underline{\Sigma}_{2|1})$$

Note that both the prior and posterior distributions are lognormal.

Some properties of the multivariate lognormal distribution are:

$$z(i|\xi)_{0.5} = \exp[\mu(i)] \tag{2.13(a)}$$

$$\langle z(i)|\xi\rangle = \exp[\mu(i)+\tfrac{1}{2}\sigma(i)] \tag{2.13(b)}$$

$$\overset{v}{\langle} z(i)|\xi\rangle = \exp[2\mu(i)+\sigma(i)](\exp[\sigma(i)]-1)$$

$$= \langle z(i)|\xi\rangle^2(\exp[\sigma(i)]-1) \tag{2.13(c)}$$

$$(i=1,2,\ldots,N).$$

$$z(q+1|q)_{0.5} = \exp[\mu(q+1)+\underline{\Sigma}_{q+1|q}\underline{\Sigma}_{11}^{-1}(\ln(\underline{z}_1)-\underline{\mu}_1)]$$

$$= \exp(\underline{\mu}_{q+1|q}) \qquad (2.14(a))$$

$$<z(q+1)|z(q),\ldots,z(1),\xi> = \exp(\underline{\mu}_{q+1|q}+\tfrac{1}{2}\underline{\Sigma}_{q+1|q}\underline{\Sigma}_{11}^{-1}\underline{\Sigma}'_{q+1|q})$$

$$= \exp(\underline{\mu}_{q+1|q}+\tfrac{1}{2}\sigma_{q+1|q}) \qquad (2.14(b))$$

$${}^{V}<z(q+1)|z(q),\ldots,z(1),\xi> = <z(q+1)|z(q),\ldots,z(1),\xi>^2 \cdot$$

$$(\exp[\sigma_{q+1|q}]-1) \qquad (2.14(c))$$

$$(q = 1,2,\ldots,N-1).$$

Note that the posterior median and mean reduce to the form

$$a_0 z(1)^{a_1} z(2)^{a_2} \ldots z(q)^{a_q} \qquad (2.15)$$

with different values for the constant a_0, respectively. The posterior variance is another constant times the square of the posterior mean. In this case the posterior variance is dependent on the revealed path although the coefficient of variation for the posterior distributions is not. Finally, the cumulative for the lognormal is very closely related to the normal cumulative as follows:

$$F_{ln}(z|\mu,\sigma) = F_n(\ln(z)|\mu,\sigma). \qquad (2.16)$$

Although the posterior moments of the multivariate lognormal distribution are more complicated and general

functions of the revealed path than were the moments for the multivariate normal distribution, the flexibility is about the same. The assessment procedure for these two distributions would be very similar. There is only one functional form for the posterior moments to take; and once this form has been specified for the posterior median, it is fixed for the posterior mean and variance.

2.4 The Multivariate Student Distribution

The N-dimensional multivariate Student distribution is characterized by an N-dimensional vector of means $\underline{\mu}$, an NxN-dimensional matrix $\underline{T}$, and a parameter n, usually called the degrees of freedom of the distribution. The density function has the form:

$$\begin{aligned} \{\underline{x}|\xi\} &= f_S^{(N)}(\underline{x}|n,\underline{\mu}\ \underline{T}) \\ &= C\left(1+\frac{1}{n}(\underline{x}-\underline{\mu})'\underline{T}(\underline{x}-\underline{\mu})\right)^{-\frac{n+N}{2}} \end{aligned} \tag{2.17}$$

where $C = \Gamma\left(\frac{n+N}{2}\right)|\underline{T}|^{\frac{1}{2}}\Gamma\left(\frac{n}{2}\right)^{-1}(n\pi)^{-\frac{N}{2}}$. Note that the matrix $\underline{T}$ is very similar to the precision matrix of the multivariate normal distribution. For $n>2$ the expected value of $\underline{x}$ and the covariance matrix of $\underline{x}$ can be shown to be $\underline{\mu}$ and $\left(\frac{n}{n-2}\right)\underline{T}^{-1}$, respectively. Thus the vector $\underline{\mu}$ is made up of the expected values of the prior distributions of

$x(i)$ $(i=1,2,\ldots,N)$. If $\underline{W}=\underline{T}^{-1}$, then the prior variance of $x(i)$ is $\frac{n}{n-2}\, w_{ii}$ where w_{ii} is the element in the ith row and the ith column of $\underline{W}$.

In order to examine the posterior distributions of the multivariate Student distribution, it will again be helpful to partition $\underline{x}, \underline{\mu}, \underline{T}$, and $\underline{W}$. Let

$$\underline{x} = (x(1),\ldots,x(q)\,|\,x(q+1),\ldots,x(N))' = (\underline{x}_1', \underline{x}_2')',$$

$$\underline{\mu} = (\mu(1),\ldots,\mu(q)\,|\,\mu(q+1),\ldots,\mu(N))' = (\underline{\mu}_1', \underline{\mu}_2')', \tag{2.18}$$

$$\underline{T} = \begin{bmatrix} \underline{T}_{11} & \underline{T}_{12} \\ \underline{T}_{21} & \underline{T}_{22} \end{bmatrix} \quad \text{where } \underline{T}_{11} \text{ is } q\text{x}q,$$

$$\underline{W} = \begin{bmatrix} \underline{W}_{11} & \underline{W}_{12} \\ \underline{W}_{21} & \underline{W}_{22} \end{bmatrix} \quad \text{where } \underline{W}_{11} \text{ is } q\text{x}q.$$

Note that $\underline{W}_{11} = \underline{T}_{11} - \underline{T}_{12}\underline{T}_{22}^{-1}\underline{T}_{21}$. The posterior distribution $\underline{x}_2$ given $\underline{x}_1$ can be shown to equal

$$\{\underline{x}_2 \,|\, \underline{x}_1, \xi\} = F_S^{(N-q)}(\underline{x}_2 \,|\, n+q, \underline{\mu}_{2|1}, \underline{T}_{2|1}) \tag{2.19}$$

where $\underline{\mu}_{2|1} = \underline{\mu}_2 - \underline{T}_{22}^{-1}\underline{T}_{21}(\underline{x}_1 - \underline{\mu}_1)$

$$= \underline{\mu}_2 + \underline{W}_{21}\underline{W}_{11}^{-1}(\underline{x}_1 - \underline{\mu}_1)$$

and

$$\underline{T}_{2|1} = \underline{T}_{22} \frac{n+q}{n+(\underline{x}_1-\underline{\mu}_1)'(\underline{T}_{11}-\underline{T}_{12}\underline{T}_{22}^{-1}\underline{T}_{21})(\underline{x}_1-\underline{\mu}_1)} .$$

See Appendix C for a proof of this. Note that this result agrees with [10] but contradicts [25]. Also note the similarity between this posterior mean and the posterior mean for the multivariate normal. The posterior covariance matrix is

$$\frac{n+(\underline{x}_1-\underline{\mu}_1)'(\underline{T}_{11}-\underline{T}_{12}\underline{T}_{22}^{-1}\underline{T}_{21})(\underline{x}_1-\underline{\mu}_1)}{n+q-2} \underline{T}_{22}^{-1} = \tag{2.20}$$

$$\frac{n+(\underline{x}_1-\underline{\mu}_1)'\,\underline{W}_{11}^{-1}(\underline{x}_1-\underline{\mu}_1)}{n+q-2} (\underline{W}_{22}-\underline{W}_{21}\underline{W}_{11}^{-1}\underline{W}_{12})$$

which is a quadratic function of the known values of the drv with its minimum at $\underline{x}_1 = \langle \underline{x}_1 | \xi \rangle$. Note that the effect of the given values of the drv $\underline{x}_1$ on the posterior covariance matrix is a decreasing function of n, the degrees of freedom. In fact this value of n can be used to adjust the curvature of this quadratic function as shown in Figure 2.1. However, since there is only one parameter n, we cannot adjust the curvature of the quadratic functions in the different periodic posterior variance expressions independently. Once one curvature is adjusted, all of the others are also determined.

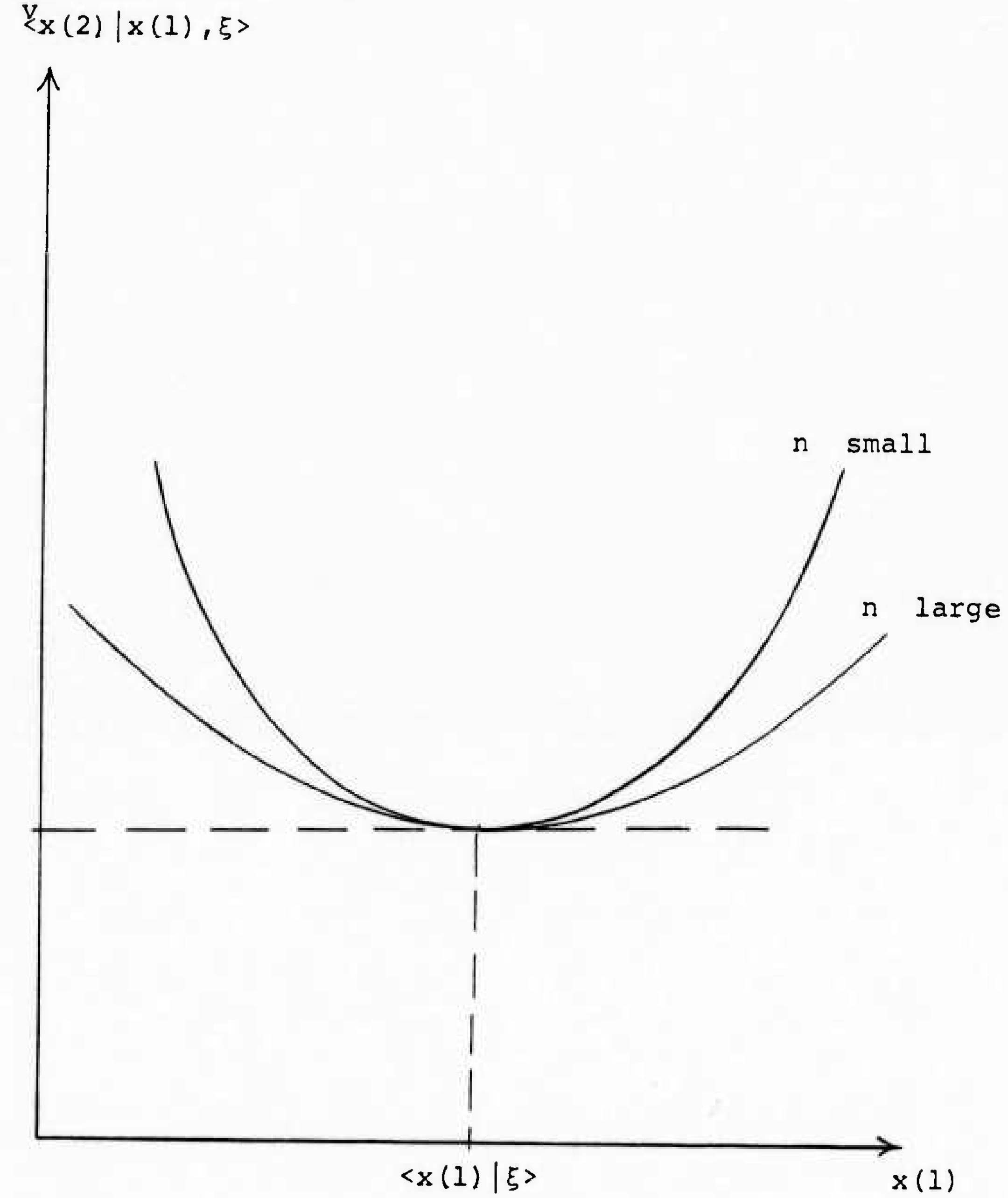


Figure 2.1. The effect of the degree of freedom parameter n on the posterior variance of x(2) given x(1)

The implicit assumption of the multivariate normal distribution is inherent in the multivariate Student distribution also. That is, the interaction between the updated probability distribution on x(i) and the given path, x(1),x(2),...,x(i-1), is contained in the off-diagonal elements of $\underline{T}$. These off-diagonal elements are closely related to the correlation coefficients.

The flexibility of the multivariate Student distribution is greater than that of the normal or lognormal distributions because it has one additional parameter. However, this increase is small in comparison to the amount of flexibility we would like. The functional forms for the posterior mean and variance in terms of the revealed path are still fixed. However, we can now make a small adjustment to all of the posterior variances, independently of the posterior means. But we would also like to have some choice over the functional forms as well.

CHAPTER 3

THE METHOD FOR UPDATING THE PRIOR DISTRIBUTIONS

3.1 Introduction

In this chapter the method for approximating the decision maker's joint probability distribution on an evolving dynamic uncertain process is developed. Ideally this method will have the following attributes: the flexibility to produce a very close approximation of the decision maker's joint probability distribution on a drv and an easily assessible form. Conceptually, our goal is to model how the decision maker updates his prior distribution on $x(i)$ to yield his posterior distribution given the values of $x(1), x(2), ..., x(i-1)$ $(i=2,3,...N)$. In other words we want to model how the decision maker uses th values of $x(1), x(2), ..., x(i-1)$ to transform his prior distribution on $x(i)$ into his posterior distribution on $x(i)$. This is illustrated in Figure 3.1. This method is basically a two point updating rodel that has enough generality to be applied in several possible ways. An example of one other application is discussed in the next section. If this process can be successfully modeled, then only N prior distributions and a reasonable number of posterior distributions must be encoded.

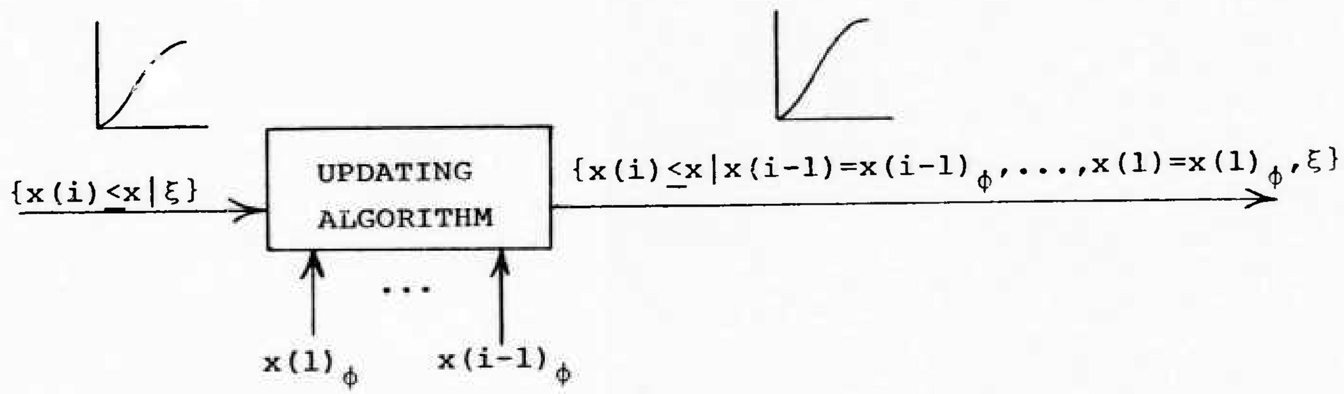


Figure 3.1. The method of updating the prior distributions

A general description of the method developed in this chapter is the following: First, the N prior probability distributions, one for the value of the drv in each time period, are assessed. Then these prior distributions are approximated by some family of continuous probability distributions, such as the normal distribution. An updating algorithm is then specified for calculating the posterior distribution of the drv in period i given the values of the drv in periods one through i-1. These posterior probability distributions are assumed to be in the same general family as the prior distributions. Additional information beyond the N prior distributions will be needed from the expert to specify and parameterize this updating algorithm for the drv. This information will take the form of assessed posterior probability distributions.

3.2 The Method

The basis of this method is a model of the decision maker's algorithm for updating his prior distribution on $x(i)$, $\{x(i)|\xi\}$, into his posterior, $\{x(i)|x(i-1),\ldots,x(1),\xi\}$ for any revealed path; $x(1)_\phi, x(2)_\phi, \ldots, x(i-1)_\phi$. Since all of our assessment techniques yield cumulative distributions, this method is based on updating the decision maker's prior cumulative to his posterior cumulative. For a discussion of probability assessment see [28]. Once the N priors and an

updating algorithm for each time period are obtained the N probability distributions that uniquely determine the joint distributions can be specified. These N distributions are:

$$\{x(1) | \xi\}$$

$$\{x(2) | x(1), \xi\}$$

$$\cdots$$

$$\{x(i) | x(i-1), \ldots, x(1), \xi\}$$

$$\cdots$$

$$\{x(N) | x(N-1), \ldots, x(1), \xi\}.$$

There are two major advantages to this approach. The first is that these are the N distributions needed for the decision analysis. Since our success will be determined by how well these distributions are approximated, we should concentrate our efforts on them. Second, the joint distribution will usually be much more intractable mathematically than will the above N distributions. So these N distributions should be easier to approximate.

However, this method is still not practical. First, it is not practical to develop separate models of the decision maker's updating process for periods two through N. This would require too many assessments of the decision maker and too much analytic effort of the analyst. For decision analyses that have been adequately modeled we have found that a recursive model of the updating algorithm provides a decent approximation. That is, a model can be specified for period i in general. There are cases where a subset of the parameters of this recursive model are functions of the period i. This complication can be handled easily though and will be discussed in the next chapter. More will be said later about this simplification and what can be done to achieve it.

Designing a model for updating an entire prior distribution into a posterior is also too difficult to accomplish practically. For this reason prior and posterior distributions are approximated by members of the same family of named distributions. For the remainder of this section all of the decision maker's prior and posterior distributions are assumed to be normal distributions. The normal family has mathematical, practical, and explanatory advantages. The ability of this method to be used with other families of named distributions is explored in the next section. The advantage of this assumption is that a named distribution is

completely characterized by two parameters. Therefore only two points of the decision maker's updating process must be modeled. The parameters of a named distribution can be calculated from any two points on its cumulative. This is indeed a great savings as well as a great simplification over specifying the updating process for every point of the distribution.

A final simplification is necessary before this method is presented in detail. This simplification is motivated by the assumption that the generating process of the drv is nonstationary as well as by the limitations on human probability assessments. A major characteristic of nonstationary processes is that the more recent the information, the more valuable it is for making inferences about that process. Therefore, when the decision maker is updating his prior distribution on $x(i)$ with the revealed path $x(1)_\phi, x(2)_\phi, \ldots, x(i-1)_\phi$, the value of $x(i-1)$ will be the most influential, the value of $x(i-2)$ the second most influential, etc. In fact, there will be some period, $i-\ell-1$, in which the value will not really affect his assessment of this posterior distribution given the more recent information $x(i-\ell)_\phi, \ldots, x(i-1)_\phi$. The implication of this is the following:

$$\{x(i)|x(i-1),\ldots,x(i-\ell),\ldots,x(1),\xi\} = \{x(i)|x(i-1),\ldots,x(i-\ell),\xi\}. \quad (3.1)$$

This type of probabilistic independence will be called <u>conditional independence</u>, and the <u>order of conditional independence</u> will then be the value ℓ for which this equation holds. Typically the order of conditional indepencence will be independent of the value of i for a given drv. It has ranged between one and three for the assessments we have done. The impact of this concept is that the updating model for period i will now only have to be a function of $x(i-1)_\phi,\ldots,x(i-\ell)_\phi$ rather than $x(i-1)_\phi,\ldots,x(i-\ell)_\phi,\ldots,x(1)_\phi$.

There are still many possible ways to model the decision maker's updating process given the above set of assumptions. The method that we have found to work best is not the most obvious, and reasons for choosing the modeling approach will be presented as the discussion proceeds. The most obvious point on a cumulative to update is the median or 0.5 fractile. The median, which is equal to the mean for the normal distribution, is the measure of a distribution's central tendency that is always assessed. In fact it is the easiest point of any distribution to assess. The most natural way to model the updating of the median in period i is to model the horizontal arrow in Figure 3.2 as a

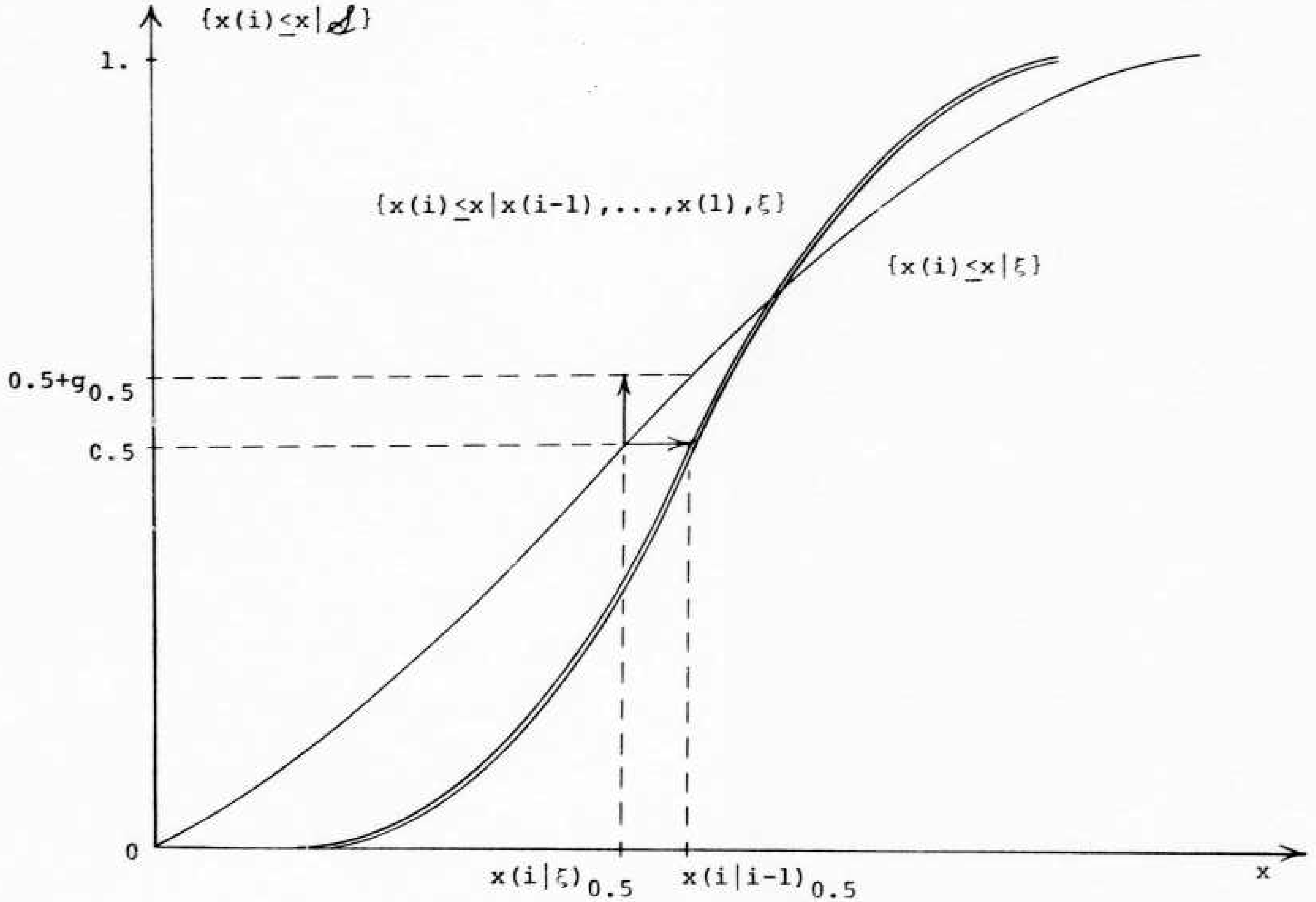


Figure 3.2. A graphical representation of the median updating algorithm

function of the revealed path, $x(i-1),\ldots,x(i-\ell)$. The single-weight curve is the prior cumulative of $x(i)$, and the double-weight curve the posterior. The horizontal arrow represents a distance and a direction that is the difference between the posterior and prior medians. It is measured in units of the drv. However, we have found that modeling in value-space (x-space), as would be done with this horizontal arrow, does not always produce the robust model we need in terms of being recursive for each period i. This is especially true of the second point to be modeled.

Rather, modeling the length of the vertical arrow in Figure 3.2 has been found to provide a very robust and accurate description of the decision maker's updating process. This vertical distance is interpreted as the delta in fractile units that must be added to or subtracted from 0.5 to find the fractile on the prior cumulative that is directly above or below, respectively, the posterior median. It can also be interpreted as the vertical distance from the posterior median to the prior cumulative, where up is the positive direction. This arrow will be referred to as $g_{0.5}$ and is in fractile space (f-space). Mathematically, this is written as:

$$x(i|i-1)_{0.5} = x(i|\xi)_{0.5+g_{0.5}}. \tag{3.2}$$

The equation for the posterior median in terms of the prior distribution and $g_{0.5}$ is:

$$x(i|i-1)_{0.5} = F_n^{-1}(0.5+g_{0.5}|x(i|\xi)_{0.5},S_{x(i|\xi)}) \quad (3.3)$$

where $S_{x(i|\xi)}$ = the prior standard deviation of $x(i)$. Solving for $g_{0.5}$

$$g_{0.5} = F_n\left[\frac{x(i|i-1)_{0.5}-x(i|\xi)_{0.5}}{S_{x(i|\xi)}}\middle|0,1\right] - 0.5. \quad (3.4)$$

That is, $g_{0.5}$ can be calculated from the posterior median, the prior median, and the prior standard deviation. One reason why this modeling strategy is successful is that the paths through time of the drv defined by a constant prior fractile (such as the path of prior medians) are generally considered more likely than other possible paths. A second reason is that this updating model is defined in terms of the decision maker's probabilistic beliefs.

The first aspect of $g_{0.5}$ to note is that it is bounded above and below by 0.5 because probabilities are restricted to lie between 0.0 and 1.0. This is a limitation in the sense that the analyst must be careful that his model observes these constraints. However, we have also found that the constraints can be useful in defining

the updating algorithm. This will be discussed in Chapter 4.

Second, $g_{0.5}$ is a function of i, $x(i-1),\ldots,x(i-\ell)$:

$$g_{0.5} = g_{0.5}(i,x(i-1),\ldots,x(i-\ell)). \tag{3.5}$$

Remember that i is an argument of $g_{0.5}$ even though it is a recursive function. That is, the functional form will rarely be a function of i, but some of the parameters may. Experience has shown that just as the median updating model constructed in x-space was not as robust as the one in f-space, having the arguments of the revealed path in f-space yields a more robust model than having them in x-space. To accomplish this $x(j)_\phi$ is transformed to its corresponding prior fractile $f(j)$ using the assessed prior distributions:

$$f(j) = \{x(j) \le x(j)_\phi | \xi\}. \tag{3.6}$$

Therefore the final functional formulation of $g_{0.5}$ is:

$$g_{0.5} = g_{0.5}(i,f(i-1),\ldots,f(i-\ell)). \tag{3.7}$$

It should be noted here that no closed-form expression exists for the normal cumulative, and therefore for its

fractiles. However the cumulative of the standardized normal distribution, $F_n(x|0,1)$, and its inverse can be calculated to great accuracy on most computers using either the "error function" library program or one of the approximations discussed in [8].

Before discussing the updating of the second point on the prior cumulative, it should be pointed out that the median updating model can be interpreted as a first approximation of the decision maker's total updating process. $g_{0.5}$ defines a horizontal translation from the prior median to the posterior median. We could in fact translate the entire prior distribution by this horizontal distance. The result would be a normal distribution with the posterior median and the prior variance. This translation is shown in Figure 3.3 as the dotted cumulative.

The approximation of the posterior distribution could be completed by updating any other fractile, say 0.9, as was done above for the median. That is, we could construct a model of $g_{0.9}$ as shown in Figure 3.4:

$$g_{0.9} = g_{0.9}(i,f(i-1,\ldots,f(i-\ell)). \tag{3.8}$$

The equation for the posterior 0.9 fractile is:

$$x(i|i-1)_{0.9} = F_n^{-1}(0.9+g_{0.9}\,|\,x(i|\xi)_{0.5},S_{x(i|\xi)}). \tag{3.9}$$

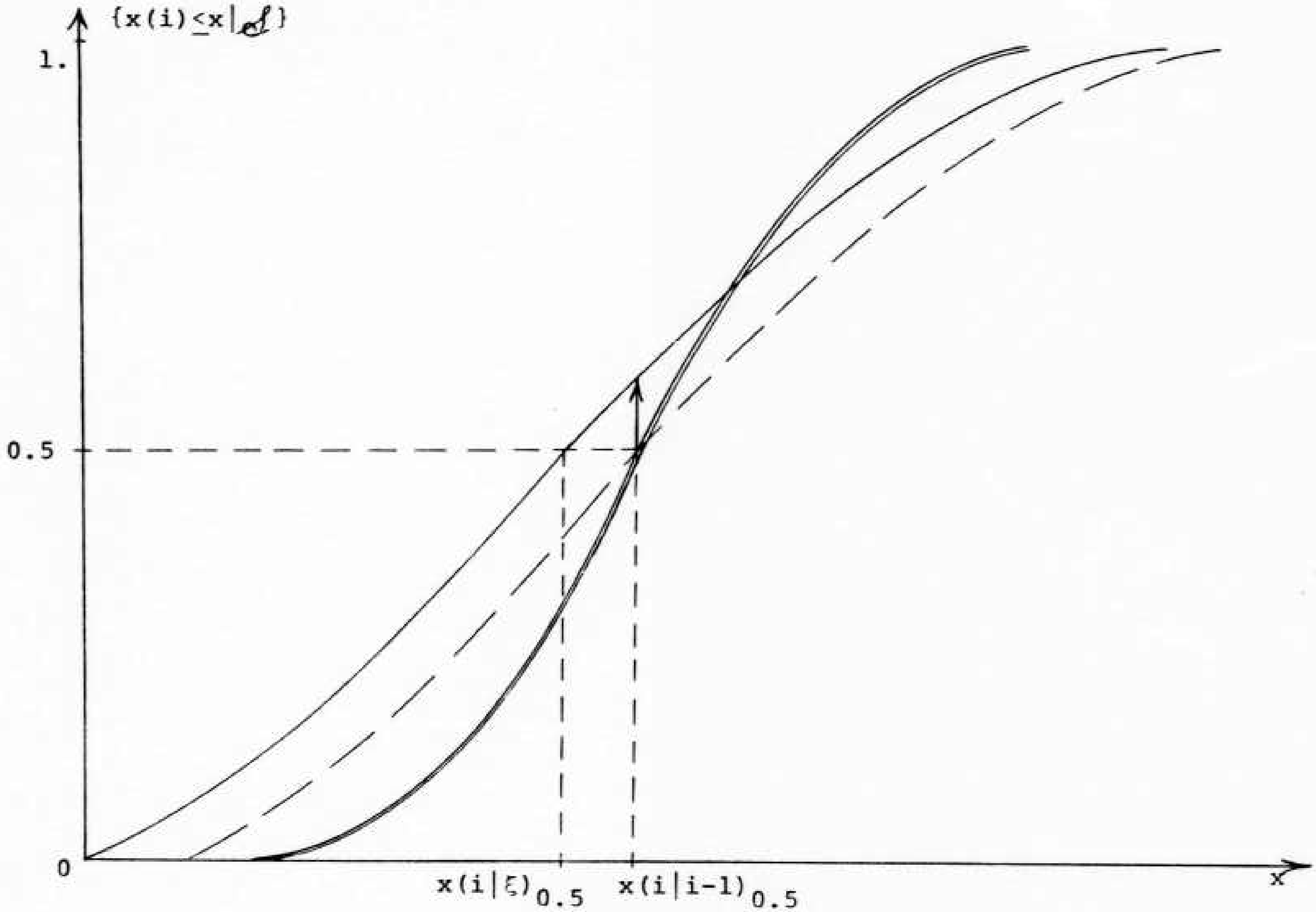


Figure 3.3. A first approximation - translation

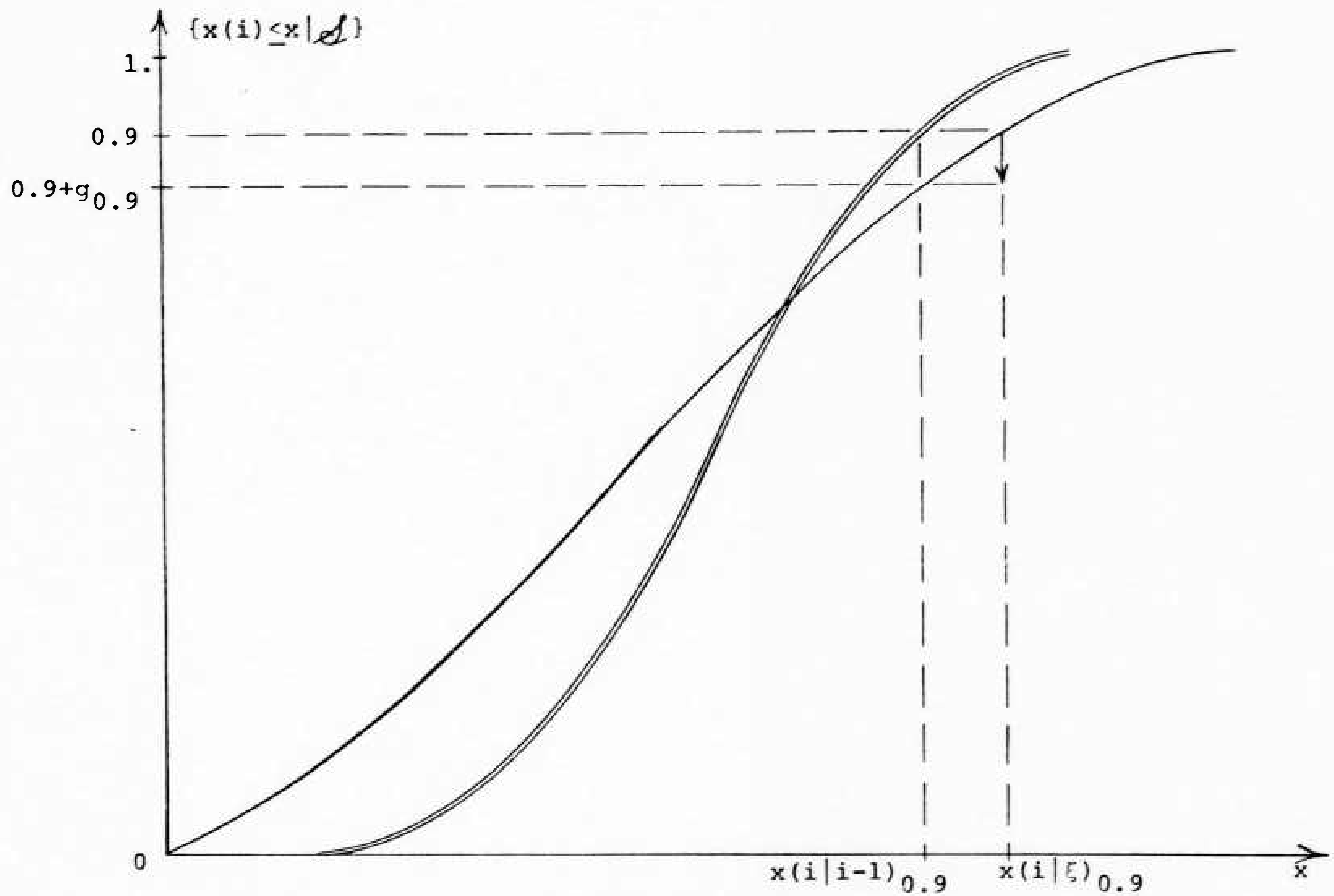


Figure 3.4. A graphical representation of $g_{0.9}$.

Solving for $g_{0.9}$

$$g_{0.9} = F_n\left[\left.\frac{x(i|i-1)_{0.9} - x(i|\xi)_{0.5}}{S_{x(i|\xi)}}\right| 0,1\right] - 0.9$$

$$= F_n\left[\left.\frac{x(i|i-1)_{0.5} + y_{0.9}S_{x(i|i-1)} - x(i|\xi)_{0.5}}{S_{x(i|\xi)}}\right| 0,1\right] - 0.9$$

where y = the standardized normal variate, $f_n(y|0,1)$

$S_{x(i|i-1)}$ = the posterior standard deviation of $x(i)$ given the values of $x(1),\ldots,x(i-1)$.

However, this approach was not successful in practice. While a decent approximation of the posterior 0.9 fractile could be obtained, the resulting approximate posterior variance was often far from the decision maker's posterior variance, which was calculated from his assessed distribution. The reason for this is that when the approximate posterior median was less (greater) than the assessed median, and the approximate posterior 0.9 fractile greater (less) than its assessed counterpart, the approximate variance was much too large (small). The result is large errors for other fractiles of the posterior distribution. (Actually we could use this approach to update every prior fractile by building a

model that not only had $i, f(i-1), \ldots, f(i-\ell)$ as arguments but f, the fractile spectrum of the posterior distribution, as well. Here the range of f would be the interval between 0.0 and 1.0. This formulation of a joint probability distribution has some interesting properties, and the results are presented in Appendix D.)

A different application of this general two-point updating approach is to a bidding situation. When a decision maker is making sequential bidding decisions with a uniform group of competitors, he would like to have a procedure for updating his uncertainty on the winning bid. In such a case the 0.9 to 0.99 fractile range of his posterior distribution is critical. This method could be used to update the 0.9 and 0.99 fractiles as discussed in the preceding paragraph. This would result in an accurate approximation of the decision maker's posterior distribution in the region of interest.

Since the median updating model defines a translation of the prior cumulative, the second updating model should be directed at rotating the prior cumulative to the standard deviation of the posterior distribution. This can be done by modifying $g_{0.9}$ slightly with the assumption that the prior and posterior medians are the same. Equation 3.10 now reduces to

$$g^*_{0.9} = F\left[\frac{y_{0.9}\, S_{x(i|i-1)}}{S_{x(i|\xi)}} \;\middle|\; 0,1\right] - 0.9. \quad (3.11)$$

The posterior standard deviation is then

$$S_{x(i|i-1)} = \frac{S_{x(i|\xi)}}{y_{0.9}} F_n^{-1}(0.9 + g^*_{0.9}|0,1)$$

$$= \frac{F_n^{-1}(0.9+g^*_{0.9}|0,1)}{F_n^{-1}(0.9\ |0,1)} S_{x(i|\xi)}. \quad (3.12)$$

Just as the median updating model could be considered a first approximation of the decision maker's updating process by translating the prior distribution to the posterior median, the standard deviation updating model can be interpreted as a second approximation. It is in fact a rotation of the prior distribution about the prior median to a distribution with the posterior standard deviation. This is shown in Figure 3.5. The dotted line represents the rotated distribution with the distance $g^*_{0.9}$ labeled. $g^*_{0.9}$ is then the vertical distance from the prior 0.9 fractile to the rotated posterior distribution with the prior median. Again, up is the positive direction.

Just as $g_{0.5}$ is bounded above and below by 0.5, $g^*_{0.9}$ also has constraints:

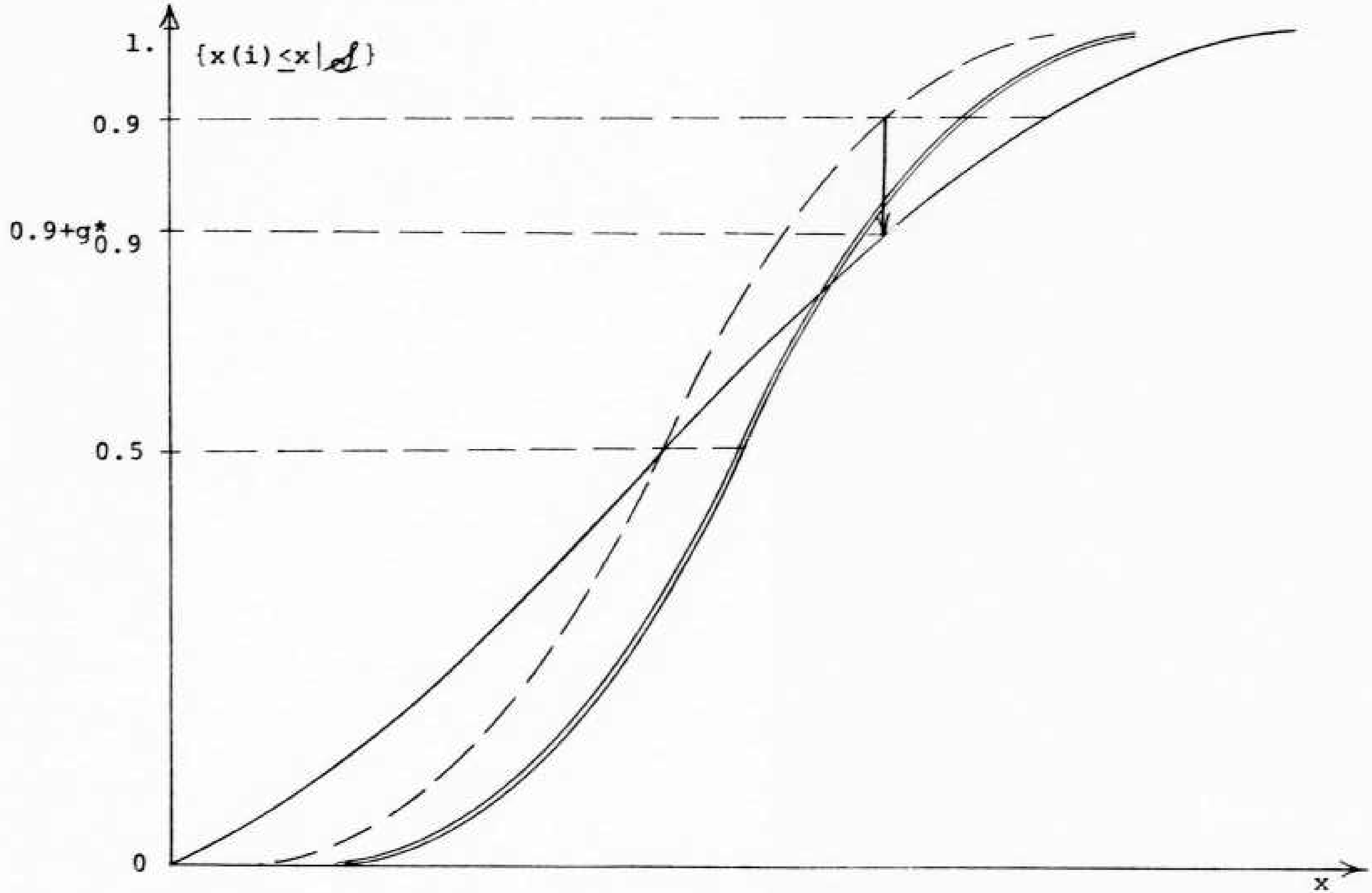


Figure 3.5. A second approximation - rotation.

$$-0.4 < g^*_{0.9} < 0.1. \qquad (3.13)$$

Actually, any fractile other than 0.5 could be chosen for the standard deviation updating model. In practice we have used the 0.7, 0.75 and 0.9 fractiles with equally satisfactory results. The constraints for $g^*_f (f>0.5)$ are:

$$f-0.5 < g^*_f < 1.0-f. \qquad (3.14)$$

It should also be noted that when the posterior variance is less (greater) than the prior variance, $g^*_{0.9}$ is less (greater) than zero.

3.3 Relaxing the Assumption of Normal Distributions

Clearly, this method would be very limited if the prior and posterior distributions could only be approximated by normal distributions. The flexibility to approximate adequately the uniform, exponential, and unimodal symmetric and skewed distributions shown in Figure 3.6 must exist if this method is to be useful. Both prior and posterior distributions may assume these shapes. Non-unimodal distributions are not considered because their occurrence indicates the need for more modeling.

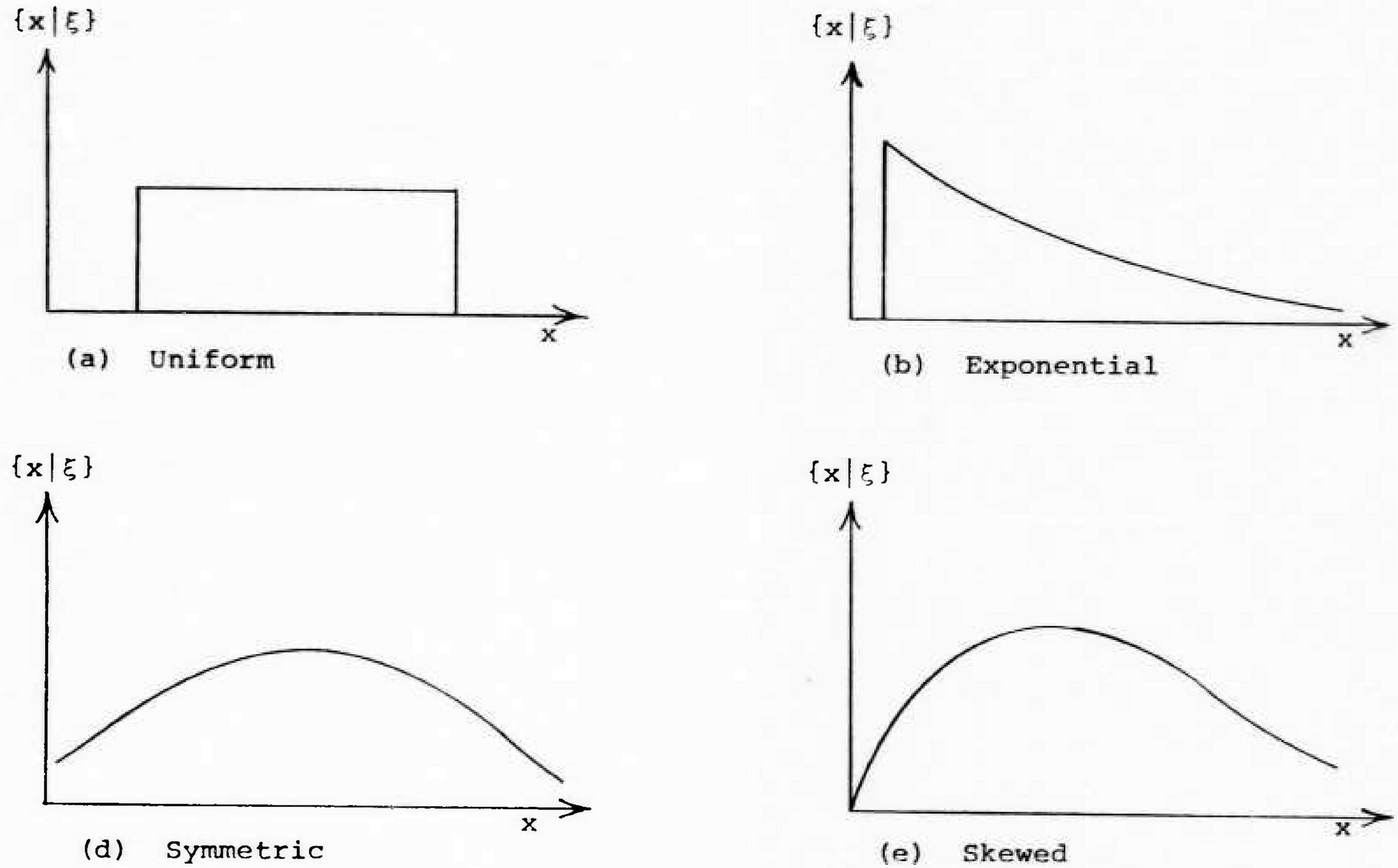


Figure 3.6. Possible shapes of assessed probability distributions

It should be pointed out here that all of the posterior distributions need not be members of the same family of named distributions. It is only necessary that all of the posterior distributions for a given period be members of the same family. However, in order for $g_{0.5}$ and $g^*_{0.9}$ to be recursive all of the prior distributions must be members of the same family. If the drv is adequately modeled as mentioned before, these restrictions should not be a major problem. It was not a serious problem for the assessments we did. In order to develop succinctly equations for the other families of named distributions, we will assume that all of the prior and posterior distributions are members of the same family for the rest of this section.

If the family of uniform distributions (Figure 3.6(a)) is the best approximation of the assessed distributions, this method can certainly be used. However, building a model of the posterior end points of the uniform distribution as a function of the revealed path would be a more direct, intuitive and rewarding approach.

Next, the exponential distribution (Figure 3.6(b)) is perhaps the easiest family to use with this method because it has only one parameter. The exponential distribution has the following density function and characteristics:

$$\{x|\xi\} = f_e(x|\lambda) = \lambda\exp[-\lambda x],\ x>0, \qquad (3.15(a))$$

$$x_{0.5} = \ln(2)(\lambda)^{-1}, \qquad (3.15(b))$$

$$\{x<x_\phi|\xi\} = F_e(x|\lambda) = 1.0-\exp(-\lambda x_\phi). \qquad (3.15(c))$$

Note that the parameter is directly related to the median.

Only the median updating algorithm of the decision maker must be modeled for this distribution. The posterior median is:

$$x(i|i-1)_{0.5} = F_e^{-1}(0.5+g_{0.5}|\lambda(i|\xi)) \qquad (3.16)$$

$$= -(\lambda(i|\xi)\ln(0.5-g_{0.5}))^{-1}.$$

Therefore $g_{0.5}$ is

$$g_{0.5} = 0.5-\exp\left[\frac{x(i|\xi)_{0.5}}{x(i|i-1)_{0.5}}\ln(0.5)\right]. \qquad (3.17)$$

For symmetric posterior distributions (Figure 3.6(c)) the normal family will often yield an acceptable approximation of the assessed distributions. Actually, many skewed distributions can be adequately approximated by the normal distribution. In fact all of our assessments were in this category.

Finally, for the skewed distribution shown in Figure 3.6(d) the lognormal family of distributions provides the most reasonable approximation for this method. The other families for approximating skewed distributions are the gamma, Weibull, Fisher, and inverse Gaussain (Wald). There are two general characteristics of a distribution that are important in maintaining a simple model of the updating algorithm. These are the complexity of the distribution's cumulative, which is the equation for the fractiles, and the relationship between the parameters of the distribution and its median and variance. Of the five families of the distributions mentioned above only two have a closed-form relationship between their median and their parameters, the lognormal and Weibull. The median of the gamma distribution can be approximated by several different functions; however, all of these functions are very cumbersome for sufficiently accurate approximations. Thus we can eliminate three families. The relationship between the variances of the remaining two distributions and their parameters are also closed-form expressions. However, the variance of the Weibull distribution,

$$\left(\frac{1}{\alpha}\right)^{\frac{2}{\beta}} \left[\Gamma\left(\frac{\beta+2}{\beta}\right) - \left(\Gamma\left(\frac{\beta+1}{\beta}\right)\right)^2 \right] \tag{3.18}$$

is a very complex function of the parameters α and β. In fact, it would require an iterative approach to calculate the parameters of this distribution given its median,

$(\ln(2)/\alpha)^{\frac{1}{\beta}}$, and variance. Therefore we are left with the lognormal distribution. Unfortunately, no closed-form expression exists for the cumulative of the lognormal distribution, just as for the normal distribution. However, the cumulative of the lognormal distribution is a function of the normal cumulative (see Chapter 2.3):

$$F_{ln}(x|\mu,\sigma) = F_n\left[\frac{\ln(x)-\mu}{\sigma} \,\middle|\, 0,1\right]. \tag{3.19}$$

Therefore the equations of the previous section for the normal family can be derived for the lognormal family. The lognormal distribution was discussed in Chapter 2.3. As a review, the following information is presented:

$$\{x|\xi\} = f_{ln}(x|\mu,\sigma) \tag{3.20(a)}$$

$$= \quad \sigma x(2\pi)^{\frac{1}{2}\ -1} \quad \exp\left[-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2\right],$$

$$x_{0.5} = \exp(\mu) \tag{3.20(b)}$$

$$\langle x|\xi\rangle = \exp(\mu+\tfrac{1}{2}\sigma) \tag{3.20(c)}$$

$$\overset{v}{\langle x|\xi\rangle} = \exp(2\mu+\sigma)(\exp(\sigma)-1.0) \tag{3.20(d)}$$

$$= \langle x|\xi\rangle^2(\exp(\sigma)-1.0).$$

For the lognormal family the equation for the posterior median is

$$x(i|i-1)_{0.5} = F_{\ln}^{-1}(0.5+g_{0.5}|\mu(i|\xi),\sigma(i|\xi)) \qquad (3.21)$$

$$= x(i|\xi)_{0.5}(\exp(\sigma(i|\xi)F_n^{-1}(0.5+g_{0.5}|0,1)))$$

where $\mu(i|\xi) = \ln(x(i|\xi)_{0.5})$

$$\sigma(i|\xi) = 2(\ln(\langle x(i)|\xi\rangle)-\mu(i|\xi))$$

$$= \ln\left[1+\frac{\check{x}(i|\xi)}{\langle x(i|\xi)\rangle^2}\right].$$

For a proof of this, see Appendix E. Solving for $g_{0.5}$ we get

$$g_{0.5} = F_n\left[\ln\left(\frac{x(i|i-1)_{0.5}}{x(i|\xi)_{0.5}}\right)\sigma(i|\xi)^{-1}|0,1\right] - 0.5. \qquad (3.22)$$

However , the solution for the posterior standard deviation must be done differently for the lognormal family. First the updated 0.9 fractile must be found under the assumption that the prior and posterior medians are equal. This assumption will be denoted by the state of information $\mathcal{S}^*$.

$$x(i|\mathcal{S}^*)_{0.9} = F_{\ln}^{-1}(0.9+g^*_{0.9}|\mu(i|\xi),\sigma(i|\xi)) \qquad (3.23)$$

$$= x(i|\xi)_{0.5}(\exp(\sigma(i|\xi)F_n^{-1}(0.9+g^*_{0.9}|0,1))).$$

Next a relationship between the f fractile of a lognormal distribution and its parameters must be used. This relationship is proven in [1].

$$x_f = \exp(\mu+\sigma F_n^{-1}(f|0,1)). \tag{3.24}$$

Therefore

$$\sigma(i|i-1) = \left[\frac{\ln(x(i|\mathcal{S}^*)_{0.9})-\mu(i|\xi)}{y_{0.9}}\right]. \tag{3.25}$$

The equation for $g^*_{0.9}$ is found by manipulating equation 3.23:

$$g^*_{0.9} = F_n\left[y_{0.9}\,\frac{\sigma(i|i-1)}{\sigma(i|\xi)}\,\middle|\,0,1\right] - 0.9. \tag{3.26}$$

The μ and σ parameters have now been updated so the posterior mean and variance can also be calculated.

3.4 Sources of Error

In this section the components of error for this approximation technique are examined. There are four components: assessment error, named distribution error,

curve-fitting error, and interpolation error. The assessment error exists because the decision maker cannot be completely accurate in the expression of his probabilistic beliefs. This research has not addressed the problem of reducing this type of error, which has been investigated elsewhere [28,29].

The second component of error results from assuming that all the posterior distributions are members of the same family of named distributions. The purpose of the previous section is to provide the flexibility for this method so that this component of error is within the acceptable region. It was not a problem in our assessments, and none of the individuals who provided the probabilistic information expressed concern over this approximation.

The third type of error results when we approximate $g_{0.5}$ and $g^*_{0.9}$ with the information contained in a number of posterior distributions. The procedure for doing this is discussed in the next chapter. Clearly, this error is a decreasing function of the number of posterior distributions we use. However, posterior distributions are not free so the analyst must make a trade-off between this error and the number of posterior distributions he assesses. We have found in doing our assessments that this error could be held to a reasonable level with 12 to 18 posterior distributions.

More posterior distributions are generally needed when the updating algorithms are functions of time or nonlinear functions of the revealed path. This error can be reduced to zero by introducing as many parameters into $g_{0.5}$ and $g^*_{0.9}$ as there are data points. However, it is often important to balance the simplicity of the model against the reduction of this curve-fitting error when the simple model will do just as well for the purpose at hand. That is, adding additional terms to $g_{0.5}$ and $g^*_{0.9}$ to reduce the curve-fitting error to zero is not justified if the fourth element of error (interpolation) is unaffected by the additions. Also, part of this curve-fitting error results from the assessment error of the decision maker.

Interpolation error results when we use posterior distributions in the decision analysis that were not assessed but calculated by the approximation technique. This error is present primarily because this specification of the decision maker's joint probability distribution is an approximation. If there were a "true" $g_{0.5}$ and $g^*_{0.9}$ and we assessed enough posterior distributions to specify them, then the interpolation error would be zero. Therefore, the only way that this error can be influenced is by having an updating algorithm that is a very close representation of the decision maker's mental process. So it is much more important to have a few terms in each of $g_{0.5}$ and $g^*_{0.9}$

[redacted]

that reflect the decision maker's updating process than to have a lot of terms that only reduce the curve-fitting error to zero.

CHAPTER 4

THE ASSESSMENT PROCEDURE FOR THE METHOD OF UPDATING THE PRIOR DISTRIBUTIONS

4.1 Introduction

The assesment procedure for the method of updating the prior distributions has five phases: two for assessment two for analysis, and a verification phase. This procedure is illustrated by a flowchart in Figure 4.1.

The decision analyst will assess the N prior distributions and gather information that will enable him to construct preliminary models of the decision maker's updating algorithm in the preliminary assessment phase. The analyst will then construct these models and approximate the prior distributions with the appropriate named distributions in the preliminary analysis phase. The final assessment phase will be devoted to the assessment of posterior distributions that will enable the analyst to complete the models of the decision maker's updating process in the final analysis phase. The last phase will be the verification of the updating models for use in the decision analysis by the decision maker. These five phases will be discussed sequentially in the next five sections of this chapter.

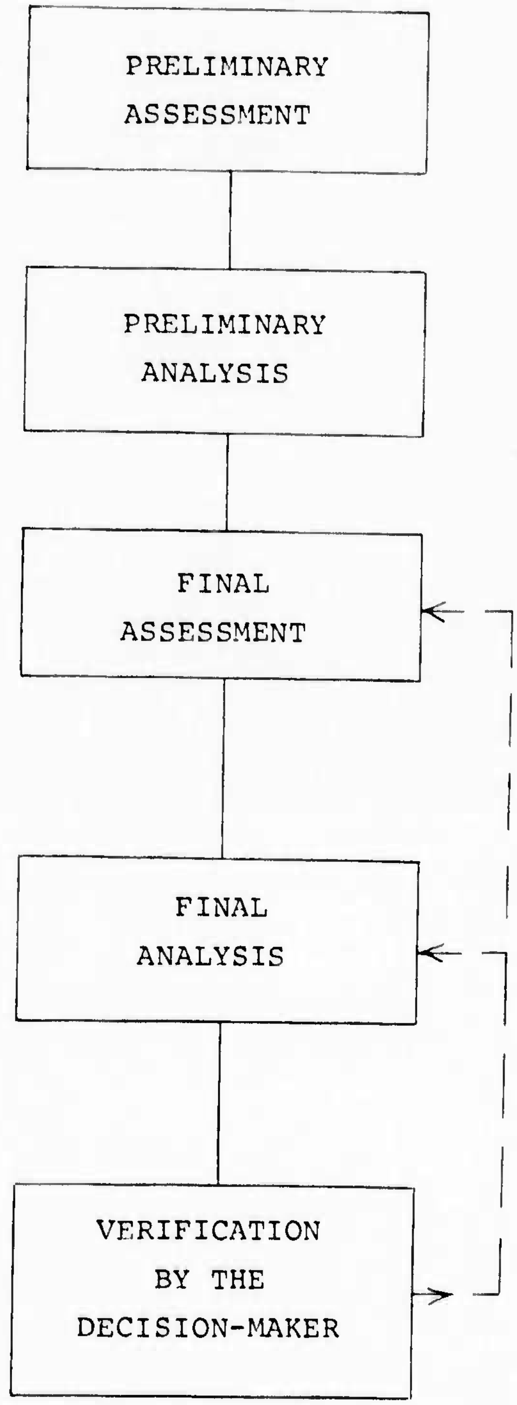


Figure 4.1. The assessment procedure

4.2 The Preliminary Assessment Phase

The preliminary assessment phase has several purposes. The first purpose is to complete the first three steps of the Spetzler and Stael von Holstein [28] encoding interview. These are the motivating, structuring, and conditioning steps. After accomplishing this initial phase by asking questions about the drv that are analgous to those for a random variable, the decision maker should be asked to quantify his thoughts in the form of N prior distributions on the drv, $\{x(i)|\xi\}$. Each of these distributions should be assessed carefully and verified afterwards because they form the foundation of this method.

Next, the analyst should obtain the information that will allow him to construct preliminary versions of $g_{0.5}$ and $g^*_{0.9}$. The first item to question the decision maker about is the order of conditional independence. The analyst should decrease the value of k, beginning at one, until the decision maker says that variations in the value of $x(N-k)$ do not affect his posterior distribution $\{x(N)|x(N-1),\ldots,x(N-k),\xi\}$. This should also be done for one or two other values of i. Usually k will be the same in each instance and $\ell=k-1$ is the order of conditional independence.

The analyst should next ask the decision maker to graph the posterior median of $x(i)$ given the value of $x(i-1)$ versus the value of $x(i-1)$, for several values of i. An example of this is shown by the solid line in Figure 4.2. To speed up the assessment process, the decision maker should not be held responsible for believing every point on the graphs. However, several points on each should be checked to make sure that he believes the general shape of the curves. The analyst should determine if the same general form is implied for each value of i. If not, then he should ask the decision maker why the graphs are different. The only reason for such a difference that we have encountered is that a future uncertain event has a major effect on the decision maker's updating process. By making the drv conditional on the outcome of this event, this problem can be alleviated. This is discussed in greater detail in Chapter 5.3.

Other useful information would be several graphs of the posterior variance of $x(i)$ given the value of $x(i-1)$ versus the value of $x(i-1)$. If the decision maker has trouble thinking in terms of the variance, then some surrogate should be used. Examples are the standard deviation or "spread of the distribution." Since this is a difficult concept to think about (even for someone trained in probability theory), the decision maker should not be held

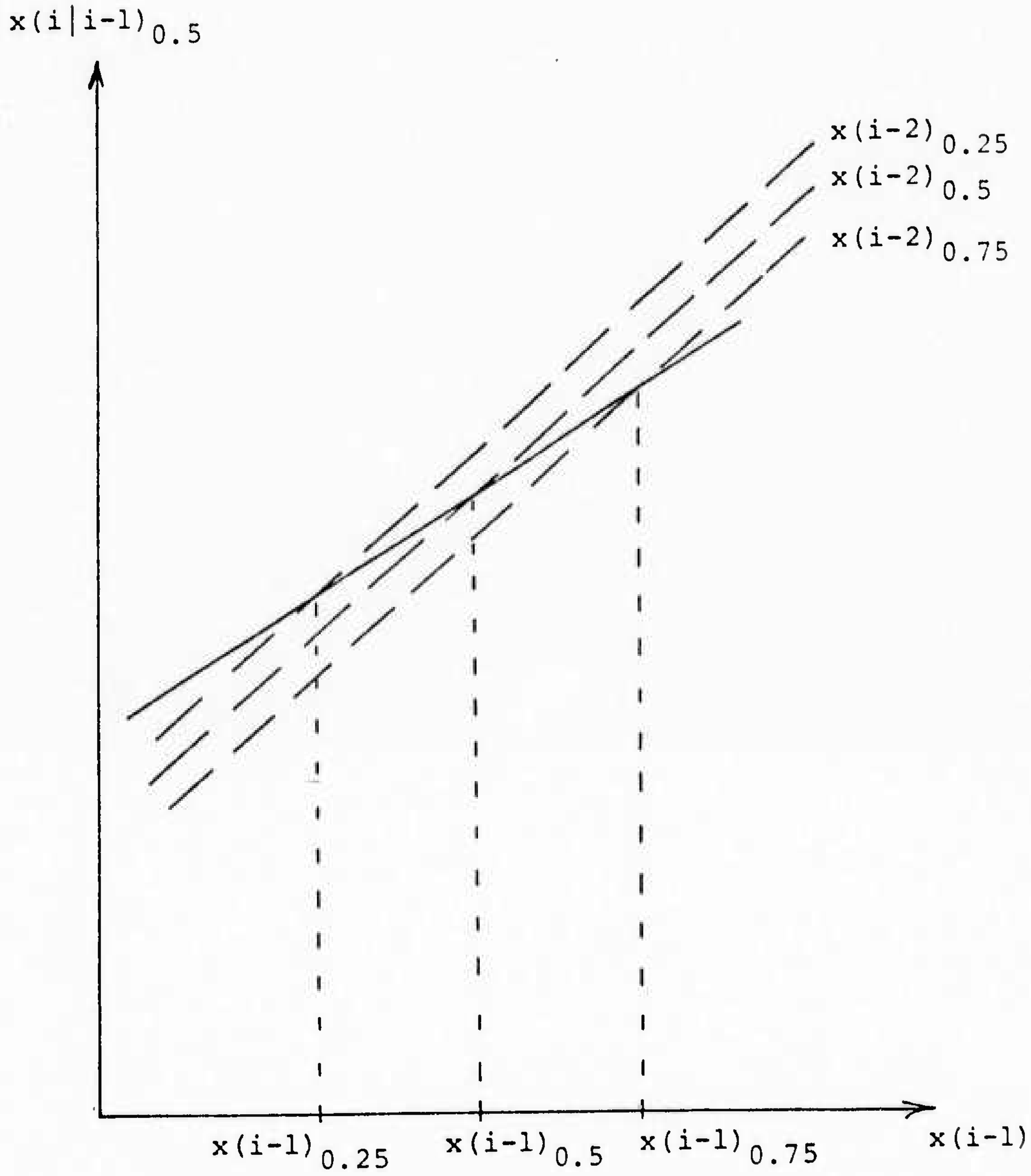


Figure 4.2. An example of a graph of $x(i|i-1)_{0.5}$ versus $x(i-1)$ and $x(i-2)$

responsible for any of the points on this curve, but only for its general shape. These graphs will be used to help determine $g^*_{0.9}$, so they should all have similar shapes.

Finally, the decision maker should be asked to describe in general terms what the effect of interactions between $x(i-1),\ldots,x(i-\ell)$ will be on the posterior distribution of $x(i)$, $\{x(i)|x(i-1),\ldots,x(i-\ell),\xi\}$. A particularly helpful method that the analyst can use is to ask the decision maker how his graphs of the posterior median $x(i|i-1)_{0.5}$ and variance $\overset{v}{x}(i|i-1)$ versus $x(i-1)$ change for several values of $x(i-2)$. The 0.25, 0.5, and 0.75 prior fractiles of $x(i-2)$ will provide sufficient information. This information can then be plotted on these original graphs as shown by the dotted lines in Figure 4.2.

A flowchart for this first phase of the assessment procedure is presented in Figure 4.3

4.3 The Preliminary Analysis Phase

There are three steps in this phase of the assessment procedure. The first is to approximate the prior distributions which were assessed in the previous phase by members of the appropriate family of named distributions. Two different methods for deriving this approximation will be described here. The first is probably not quite as accurate

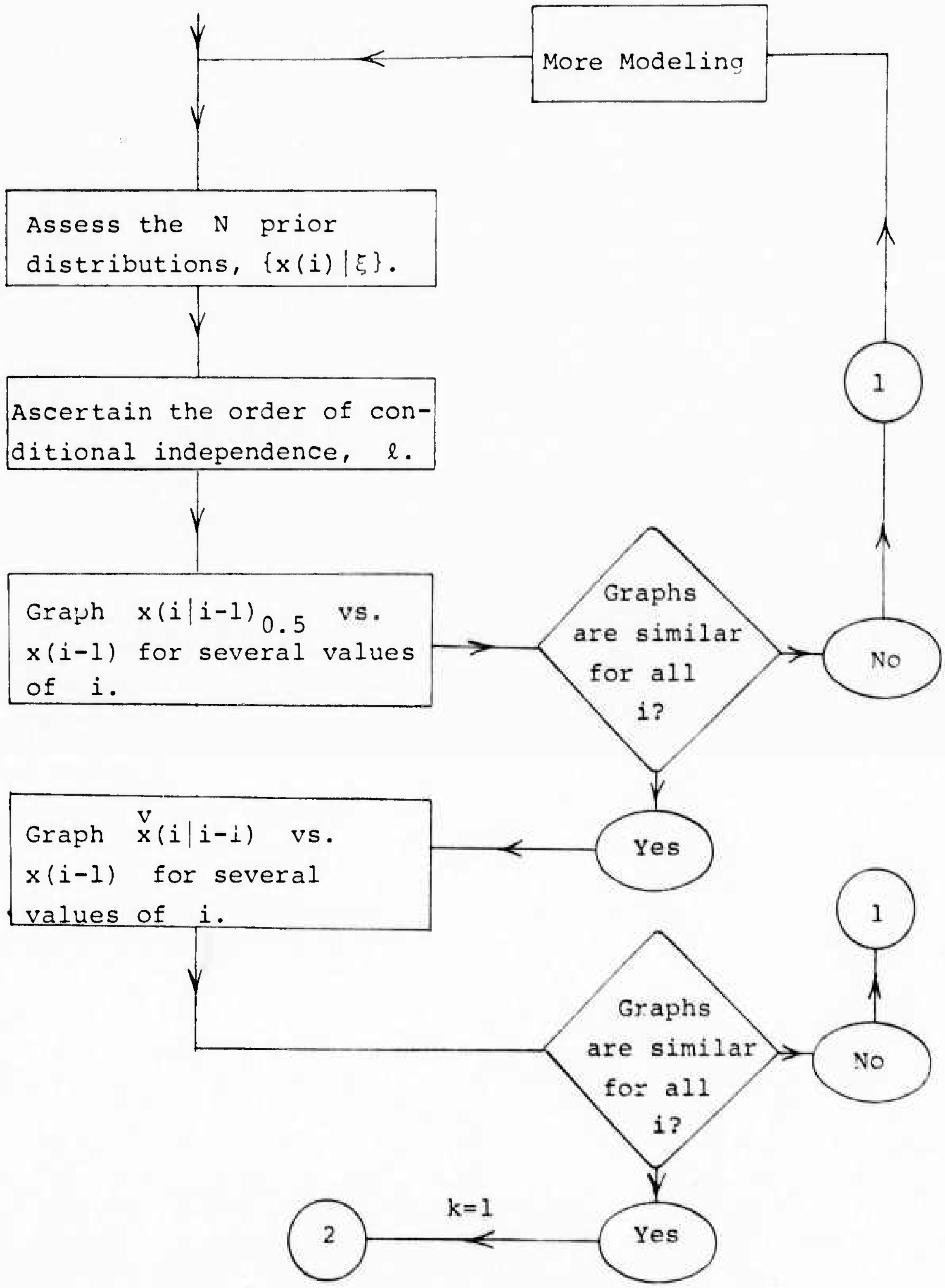


Figure 4.3. The preliminary assessment phase

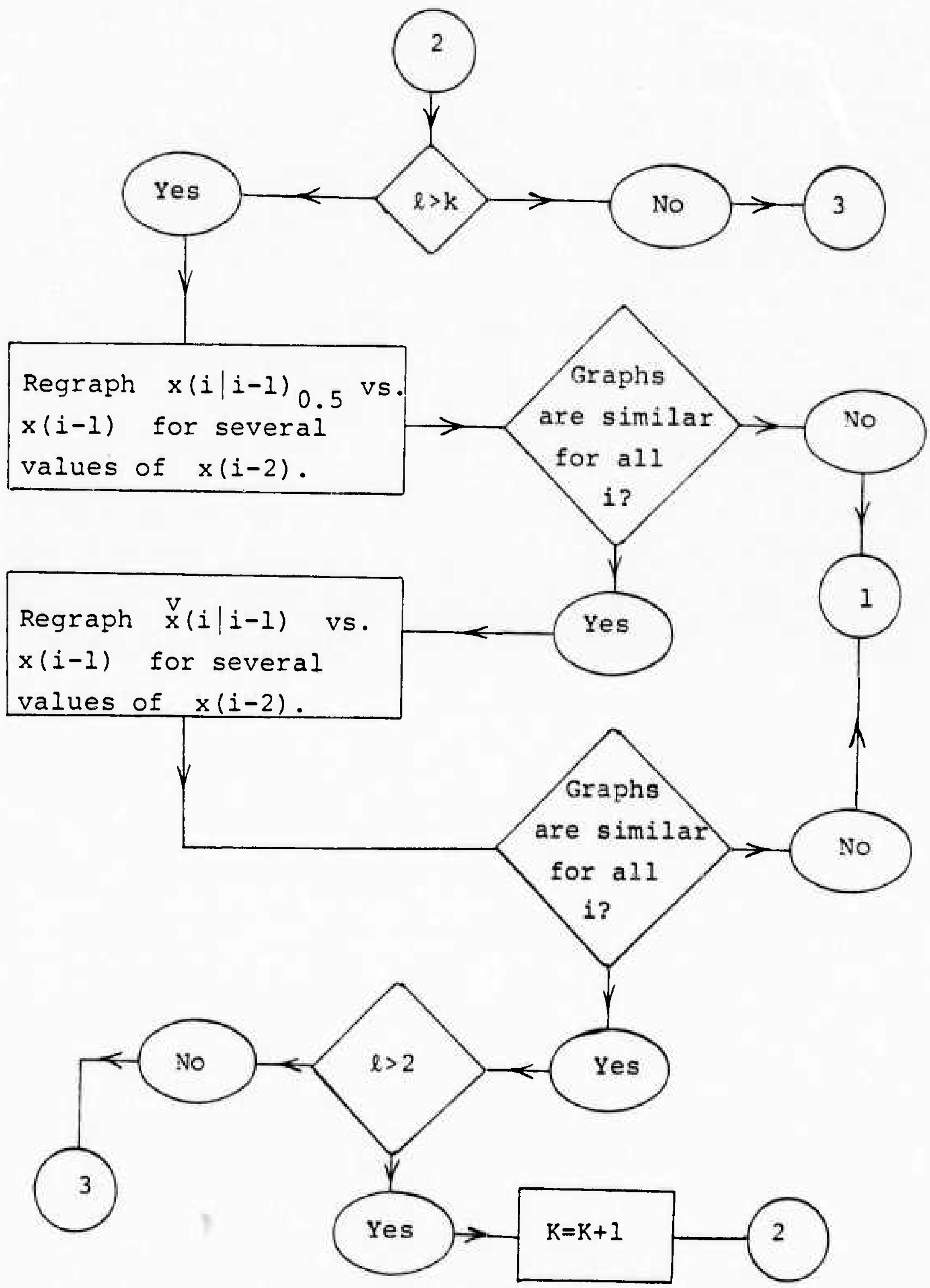


Figure 4.3. The preliminary assessment phase (cont.)

as the second but is less work. It involves estimating the parameters directly from the assessed cumulatives. For the normal distribution the two parameters correspond to the median (mean) and standard deviation. So for each prior distribution the median has been assessed. The standard deviation can be approximated by rewriting the definition of standardized normal variate:

$$S_x = \sigma = \left| \frac{x_{0.5} - x_a}{y_a} \right| \tag{4.1}$$

where $a \neq 0.5$. The average of σ calculated this way for several values of a should yield a good approximation of the prior normal standard deviation.

The exponential distribution can be approximated directly from the assessed median by using equation 3.15(b).

The prior parameters of the lognormal distribution are best approximated by the assessed median and mean. The mean can be determined graphically. See [13] for a discussion of this graphical technique. Then

$$\mu(i|\xi) = \ln(x(i|\xi)_{0.5}) \tag{4.2}$$

$$\sigma(i|\xi) = 2[\ln(\langle x(i)|\xi\rangle) - \mu(i|\xi)]. \tag{4.3}$$

However , a more refined method for calculating these parameters is least squares. See [22,26] for a discussion of least squares. For each of these distributions there is a linear equation relating the parameters and fractiles of the distribution:

$$\mu + y_f\sigma = x_f \qquad \text{(normal)} \qquad (4.4(a))$$

$$\mu + y_f\sigma = \ln(x_f) \qquad \text{(lognormal)} \qquad (4.4(b))$$

$$\lambda = -(x_f)^{-1} \ln(1.0-f). \qquad \text{(exponential)} \qquad (4.4(c))$$

Each point that the analyst assessed on a given cumulative can be translated into an equation to be used in the least squares estimation of the parameters of that prior distribution.

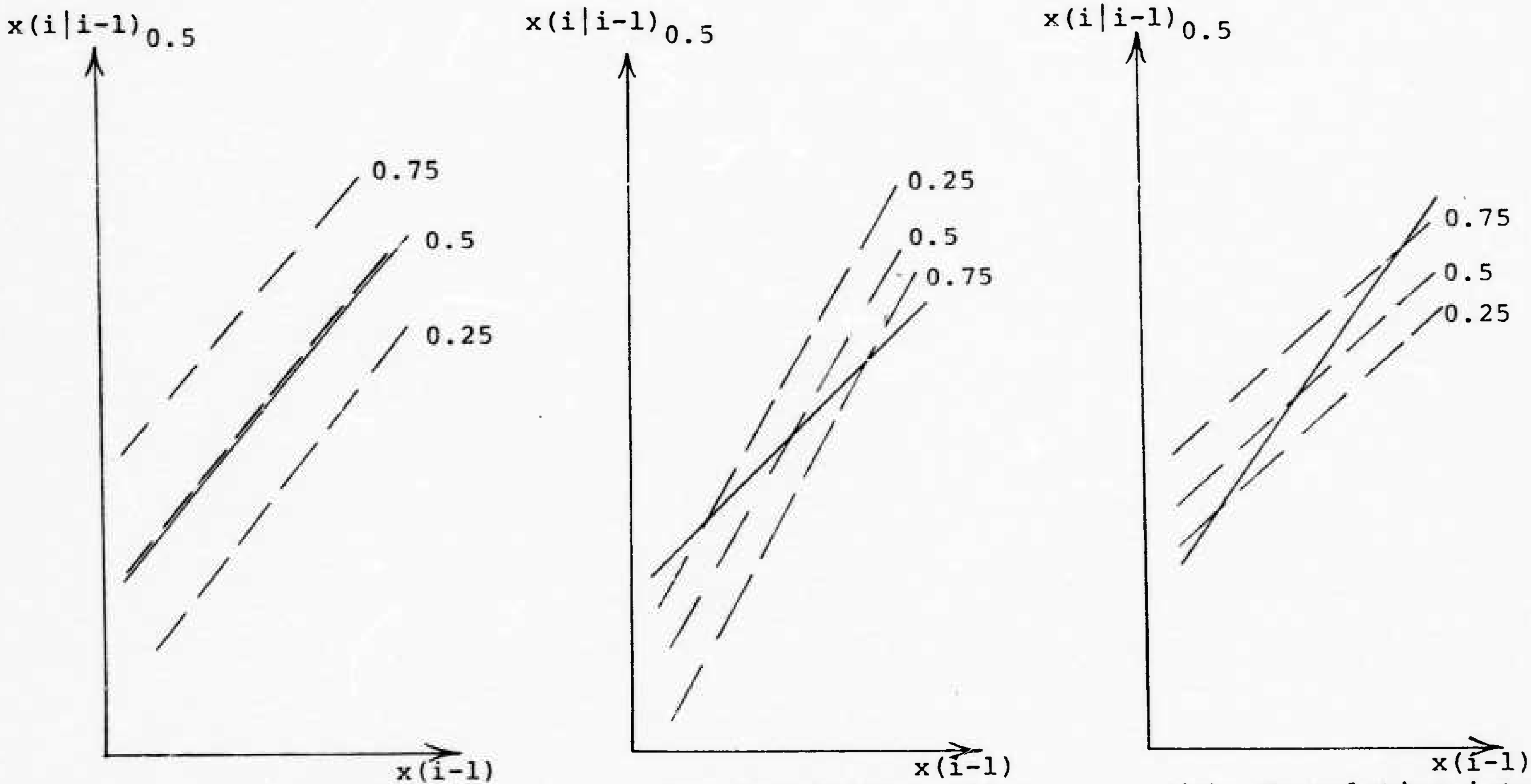
The next step is to specify preliminary forms for $g_{0.5}$ and $g^*_{0.9}$ by using the information obtained in the previous phase. For the remainder of this chapter we will assume that the order of conditional independence is two. This was the most frequent value for the assessments we did. The procedure for an order of conditional independence equal to one will be clear after this discussion; and it will be easily generalized for values greater than two. The functional form assumed for both $g_{0.5}$ and $g^*_{0.9}$ in period i is the following:

$$a_0 + h_1\,(f(i-1)) + h_2\,(f(i-1), f(i-2)). \qquad (4.5)$$

This form is completely general and complements the assessment procedure very nicely. a_0 is a constant. $h_1(f(i-1))$ is a function of only the previous value. The graphs of $x(i|i-1)_{0.5}$ and $\overset{v}{x}(i|i-1)$ versus $x(i-1)$ will be used to specify its preliminary form. $h_2(f(i-1), f(i-2))$ will then be specified by the remaining graphs obtained in the preliminary assessment phase. Note that this term will not be present when i equals two.

For nearly every assessment that we have done, in linear or piecewise linear form for $h_1(f(i-1))$ has been adequate for both $g_{0.5}$ and $g^*_{0.9}$. However, any functional form is possible as long as the constraints of these two models are met, and the analyst should be open to several possibilities in this phase of the assessment procedure.

There are numerous possibilities for $h_2(f(i-1), f(i-2))$. Some of the more common ones will be discussed here and are represented by the graphs in Figure 4.4. These are the types of graphs that would be ascertained in the last two sections of the preliminary assessment phase. Since $g_{0.5}$ and $g^*_{0.9}$ are similar in terms of functional forms, we will only discuss $g_{0.5}$ in detail. A few words will be said about $g^*_{0.9}$ later. The graph in Figure 4.4(a) implies that there is no real interaction between $f(i-1)$ and $f(i-2)$



(a) No interaction (b) Trend interaction (c) Correlation interaction

Figure 4.4 Possible types of interaction between x(i-1) and x(i-2) on the posterior median

in the term $h_2(f(i-1),f(i-2))$. This is the type of behavior implied by the multivariate named distributions and was not represented in our assessments. In this case another linear relationship

$$a_1(f(i-2)-a_2) \qquad (4.6)$$

would probably be adequate. For the case depicted in the figure a_1 would be positive and a_2 would equal 0.5. At this point it is not necessary that the analyst worry about whether the constraints of $g_{0.5}$ and $g^*_{0.9}$ are met. However, the constraints can be used to change the above linear expression into a nonlinear one in such a way that the constraints are not violated. This can best be explained with the aid of a graph. Suppose that the decision maker's graph of $x(i|i-1)_{0.5}$ versus $x(i-1)$ has been translated into a graph of $g_{0.5}$ versus $f(i-1)$ as shown in Figure 4.5 by the solid line. The beliefs shown in Figure 4.4(a) imply that if $f(i-2)$ were 0.75, then $g_{0.5}$ would be above the solid line plotted in Figure 4.5 for all values of $f(i-1)$. One way to achieve this is to multiply the above linear expression for h_2

$$a_1(f(i-2) - a_2)$$

by the distance between the constraint and the solid line,

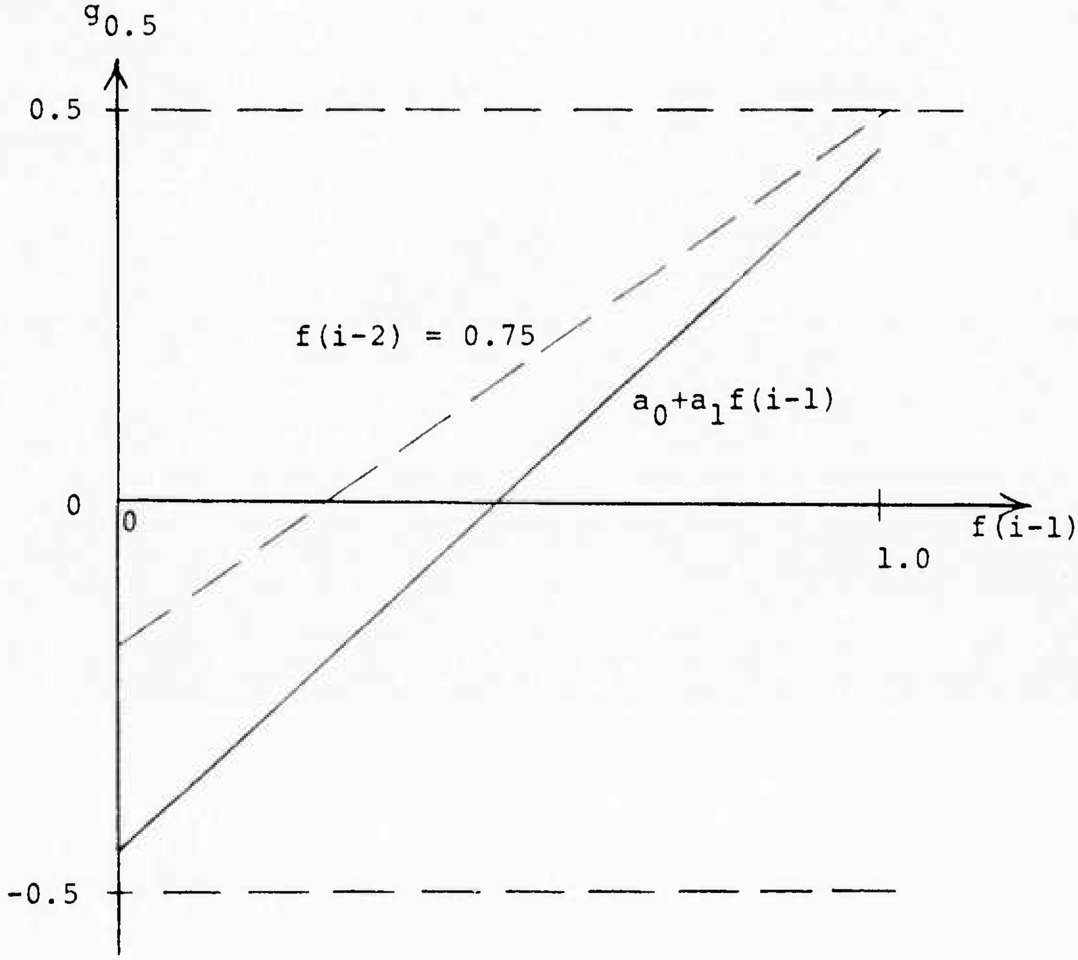


Figure 4.5. The effect of the nonlinear multiplicative factor

$$0.5-(a_3+a_4 f(i-1)). \qquad (4.7)$$

This new expression is

$$a_1(f(i-2)-a_2)(0.5-a_3-a_4 f(i-1)). \qquad (4.8)$$

The dotted line in Figure 4.5 represents $g_{0.5}$ for $f(i-2)$ equal to 0.75 and a positive value of a_1. If $f(i-2)$ were less than 0.5, the expression would be

$$a_1(f(i-2)-a_2)(a_3+a_4 f(i-1)-(-0.5)). \qquad (4.9)$$

This nonlinear expression implies the same general behavior as shown in Figure 4.4(a). However, the decision maker's assessments in the next phase must be used to determine whether it is a better characterization of his beliefs than the strictly linear relationship. Finally, the analyst may want to try nonlinear functions of $f(i-2)$ also.

The graphs in Figure 4.4(b) imply that the trend in prior fractiles is the important characteristic to be captured in $h_2(f(i-1),f(i-2))$. That is, if $f(i-2)$ is less than $f(i-1)$, then the posterior median is greater than if $f(i-2)$ equalled $f(i-1)$. The linear form of this expression is

$$a_1(f(i-1)-f(i-2))+a_2 \qquad (4.10)$$

where a_1 is positive. This belief prevailed in several of our assessments and was adequately approximated by either this linear expression or the nonlinear expression analogous to equations 4.8 and 4.9 above. Of course other nonlinear expressions can also be used.

Another popular belief reflected in our assessments was the one depicted in Figure 4.4(c). This is more like the correlation behavior of the multivariate named distributions, and the opposite of the trend behavior in Figure 4.4(b). If $f(i-2)$ is less than $f(i-1)$, then the posterior median is decreased. The linear expression for this behavior is exactly the same as equation 4.10, except that a_1 is negative.

Finally, we should discuss how the analyst can model the decision maker's belief that his posterior medians will describe a cyclic path through time. There are two ways to do this. The decision maker who already has a lot of information about this cyclic path will provide a cyclic path of prior medians in the preliminary assessment phase. This cyclic path of prior medians and any of the above functional forms for $g_{0.5}$ will result in a cyclic path of posterior medians. When the decision maker does not have much prior information about the cyclic path of the drv but will be gaining some as the future values of the drv are revealed, a different approach must be taken. In this case

the path of prior medians will have little or no cyclic shape. However, a plot of the revealed values of the drv in prior fractile units over time should reveal a cyclic shape. This cyclic shape can easily be translated into a cyclic path of posterior medians by letting $h_1(f(i-1))$ be a decreasing function of $f(i-1)$. Of course, there are also more complex ways to achieve this.

In four of our assessments the posterior standard deviation was an increasing function of $x(i-1)$ and was adequately approximated by one of the functional forms described above for the posterior median. However, a different belief, which was reflected in three of our assessments, is that the standard deviation is either an increasing or decreasing function of the distance between the actual value of $x(i-1)$ and $x(i|\xi)_{0.5}$. That is, if $x(i-1)$ equals the decision maker's prior median in period i-1, then his posterior standard deviation in period i is either a minimum or maximum. In this case, h_1 is a function of $|f(i-1)-0.5|$. A possible permutation of this belief for $\ell=2$ is that the posterior standard deviation in period i is either a minimum or maximum when $f(i-1)$ equals $f(i-2)$.

The last step is to determine which posterior distributions are to be assessed in the final assessment phase in order to find the parameters of $g_{0.5}$ and $g^*_{0.9}$. There are

no special techniques for doing this although a few rules of thumb are useful. It is a good idea to assess enough posterior distributions in at least three time periods so that the parameters of $g_{0.5}$ and $g^*_{0.9}$ can be determined completely for each of these time periods. Thus, for example, the analyst might want to assess four or five posterior distributions for i equal to two, four, and six. If h_1 and/or h_2 are expected to be nonlinear, then even more posterior distributions may be needed. Frequently one or more parameters of $g_{0.5}$ and $g^*_{0.9}$ are functions of i. For this reason it will also be useful to assess at least one or two posterior distributions for all or nearly all of the remaining time periods. This will provide the information to approximate these parameters as functions of i, if necessary. Finally, the analyst should guard against assessing too many distributions. The accuracy of the approximation should be balanced against the cost of the assessment. A flow graph of this phase of the assessment procedure is presented in Figure 4.6.

4.4 The Final Assessment Phase

The final assessment phase will hopefully only involve assessing the posterior distributions decided upon in the preliminary analysis phase. Again the assessment techniques should conform to those described in [28]. However, the analyst should be checking to make sure that the medians and

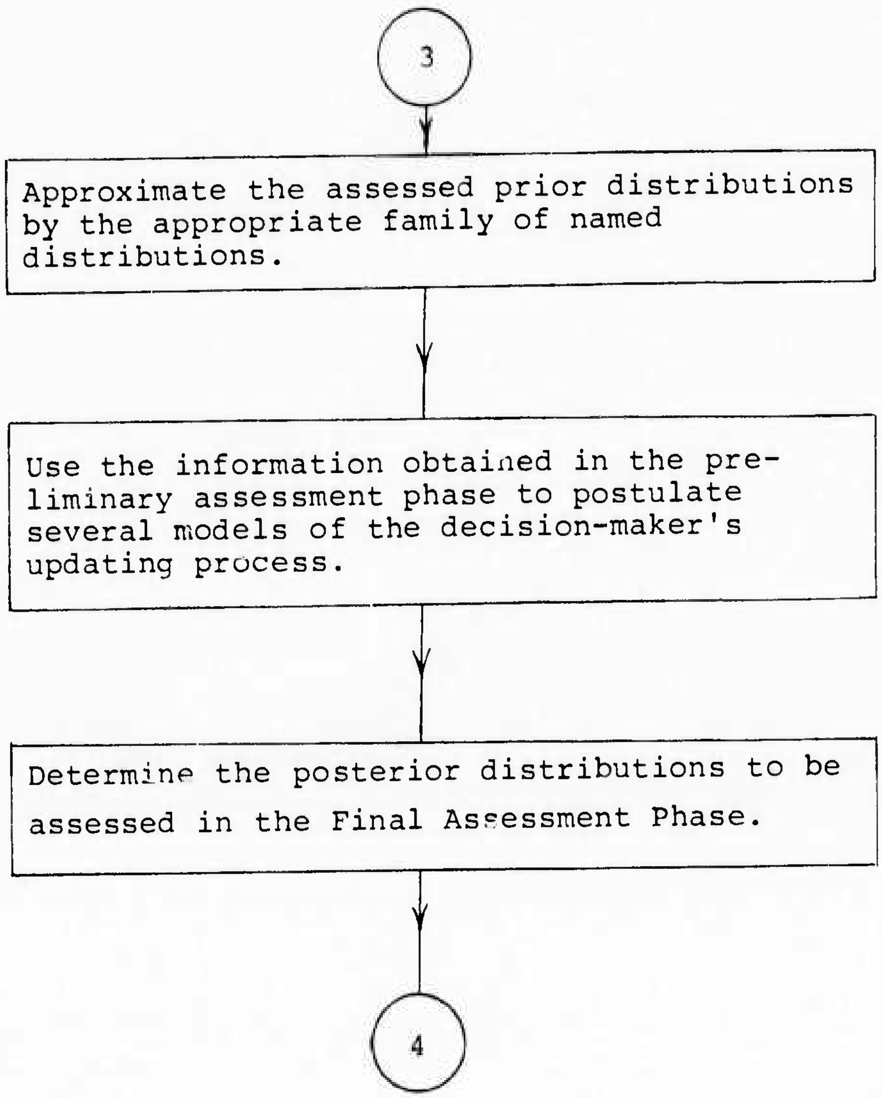


Figure 4.6. The preliminary analysis phase

variances of the assessed posterior distributions correspond generally to the behavior described by the decision-maker in the preliminary assessment phase. Only qualitative checks by the analyst are required here. If the assessed distributions do not exhibit the same behavior as the preliminary graphs, then the decision maker should be consulted and asked to decide which he believes. If he believes the assessed posterior distributions, then the analyst may want to adjust which posterior distributions he assesses. If the decision maker believes the graphs he drew in the preliminary assessment phase, then he should change the appropriate posterior assessments. This procedure is depicted in Figure 4.7.

4.5 The Final Analysis Phase

There are three basic steps in this final analysis phase. The first is to approximate the posterior distributions assessed in the previous phase by members of the appropriate family of named distributions. The same process should be used here as was used for the prior distributions. Next, $g_{0.5}$ and $g^*_{0.9}$ should be calculated for each of these posterior distributions. For the normal family the equations used for this purpose would be 3.4 and 3.11. Analogous equations can be found in Chapter 3.3 for the other families of named distributions. This information should then be used to solve for the parameters of $g_{0.5}$

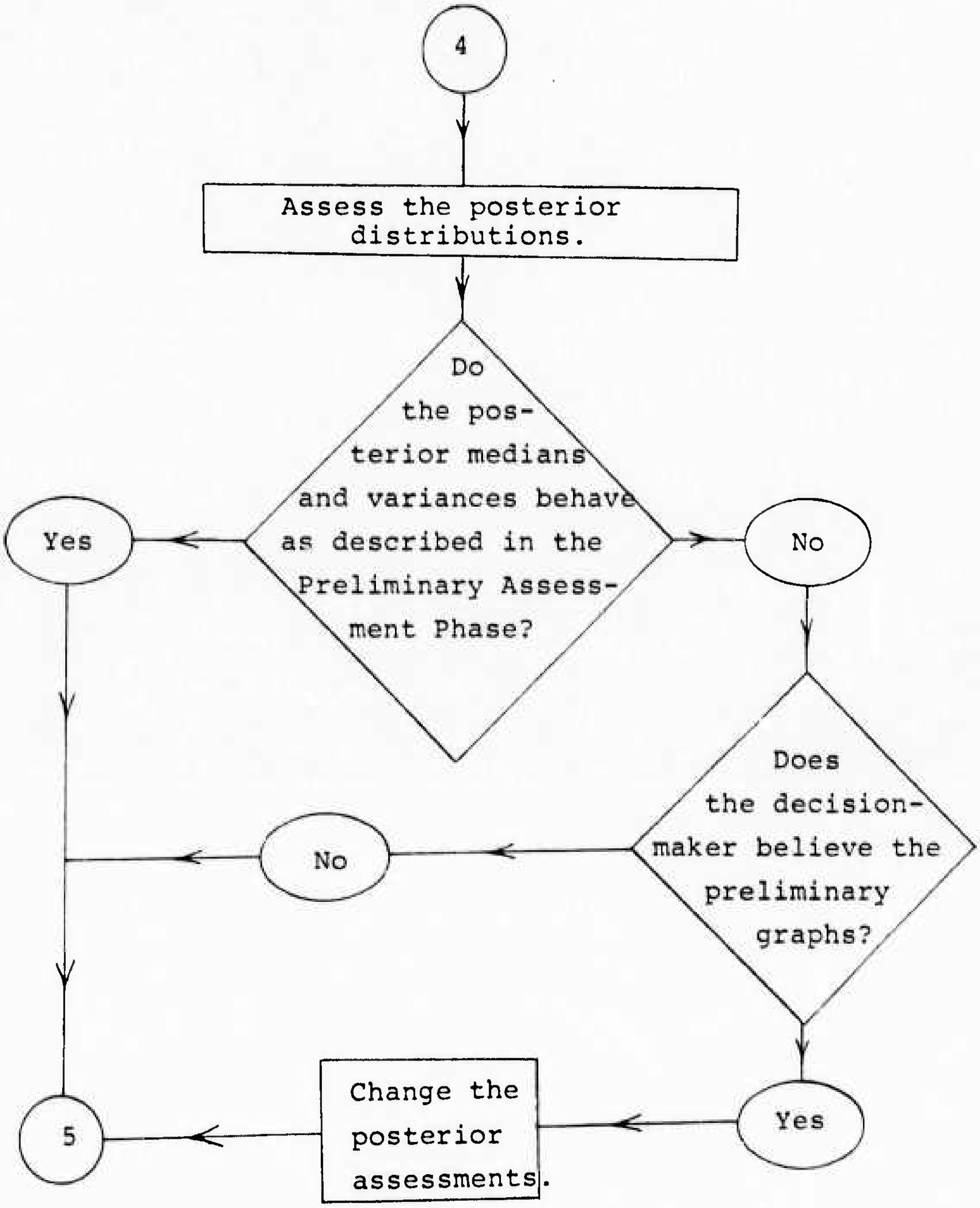


Figure 4.7. The final assessment phase

and $g^*_{0.9}$. There are two methods that can be used to find these parameters. Since all of the updating models discussed in Section 3 are linear in terms of the parameters. least squares can be used. Each assessed posterior distribution would provide a data point for the parameter of both the median and standard deviation updating models. Care should be taken however to determine whether any of the parameters should be functions of the period i. The second method is a graphical one which will be discussed in the next chapter.

At this point it would be wise for the analyst to compare the assessed posterior distributions to the approximate ones given by the prior distributions and $g_{0.5}$ and $g^*_{0.9}$. He should be convinced that the approximation is sufficiently accurate for the decision analysis. Otherwise he may want to try more complex functions for $g_{0.5}$ and $g^*_{0.9}$. Figure 4.8 presents a flowchart for this phase.

4.6 The Verification by the Decision Maker

By this time the first step in the verification procedure has been completed. That is, the deicison maker and analyst have agreed on the general forms of $g_{0.5}$ and $g^*_{0.9}$. Now all that has to be determined is whether the approximation is good enough so that the decision maker is willing to use it in the determination of his optimal decision. The

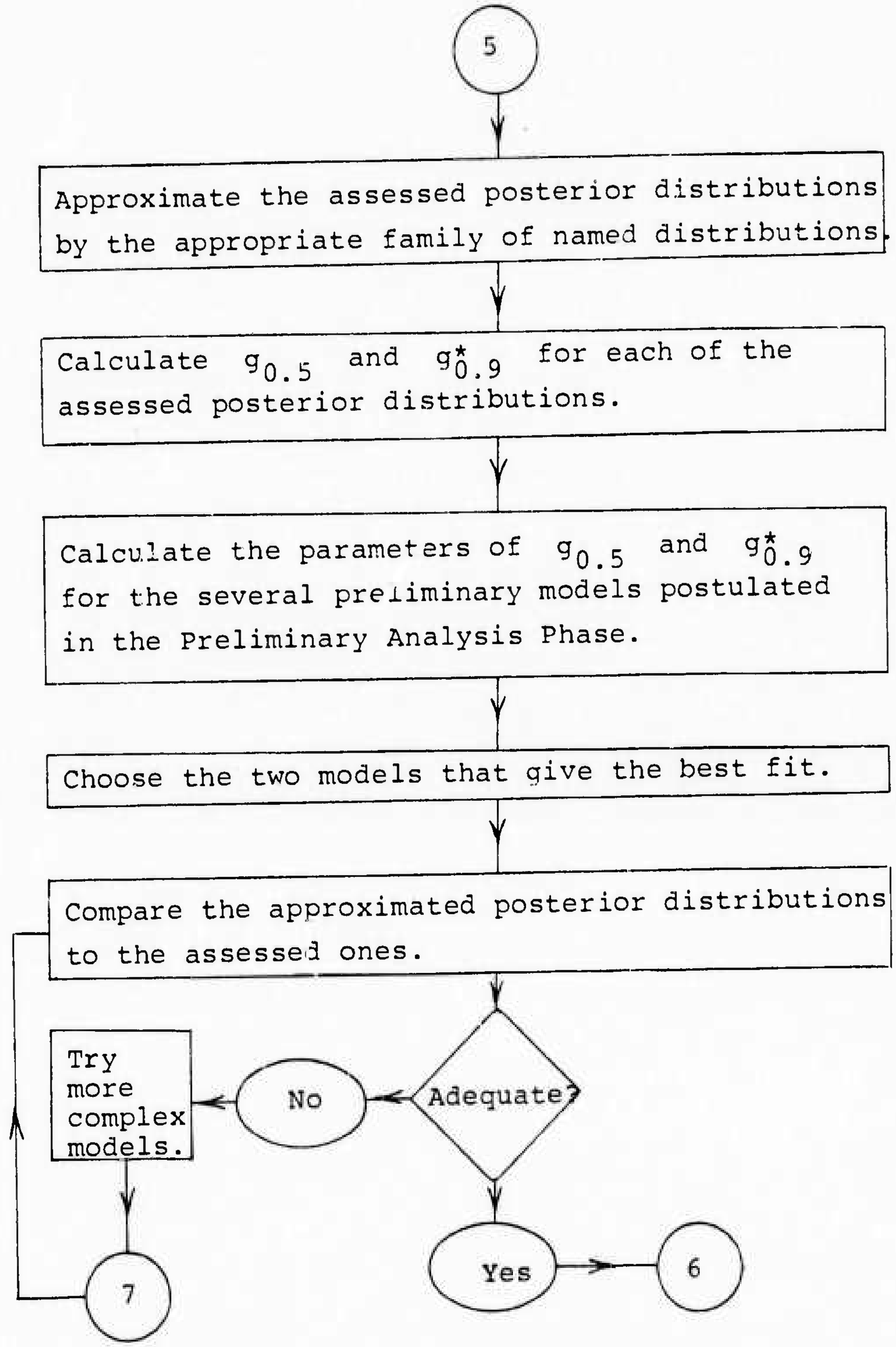


Figure 4.8. The final analysis phase

first step is to show the decision maker the comparisons between his assessments and the approximations which were computed in the final analysis. If he feels these approximations are adequate, the analyst should assess several new distributions for which he has already prepared approximations. If these approximations are also close enough, then the decision maker will most likely accept the approximation. If not, the analyst should determine what has gone wrong. Usually it will be that there is a better form for $g_{0.5}$ or $g^*_{0.9}$. If so, the analyst should repeat the final analysis. By agreeing to use this approximation of his joint distribution, the decision maker is saying that he is indifferent to accepting the actual values of $x(1), x(2), \ldots, x(N)$ or values generated by a simulation using this approximate joint distribution. This phase is depicted in Figure 4.9.

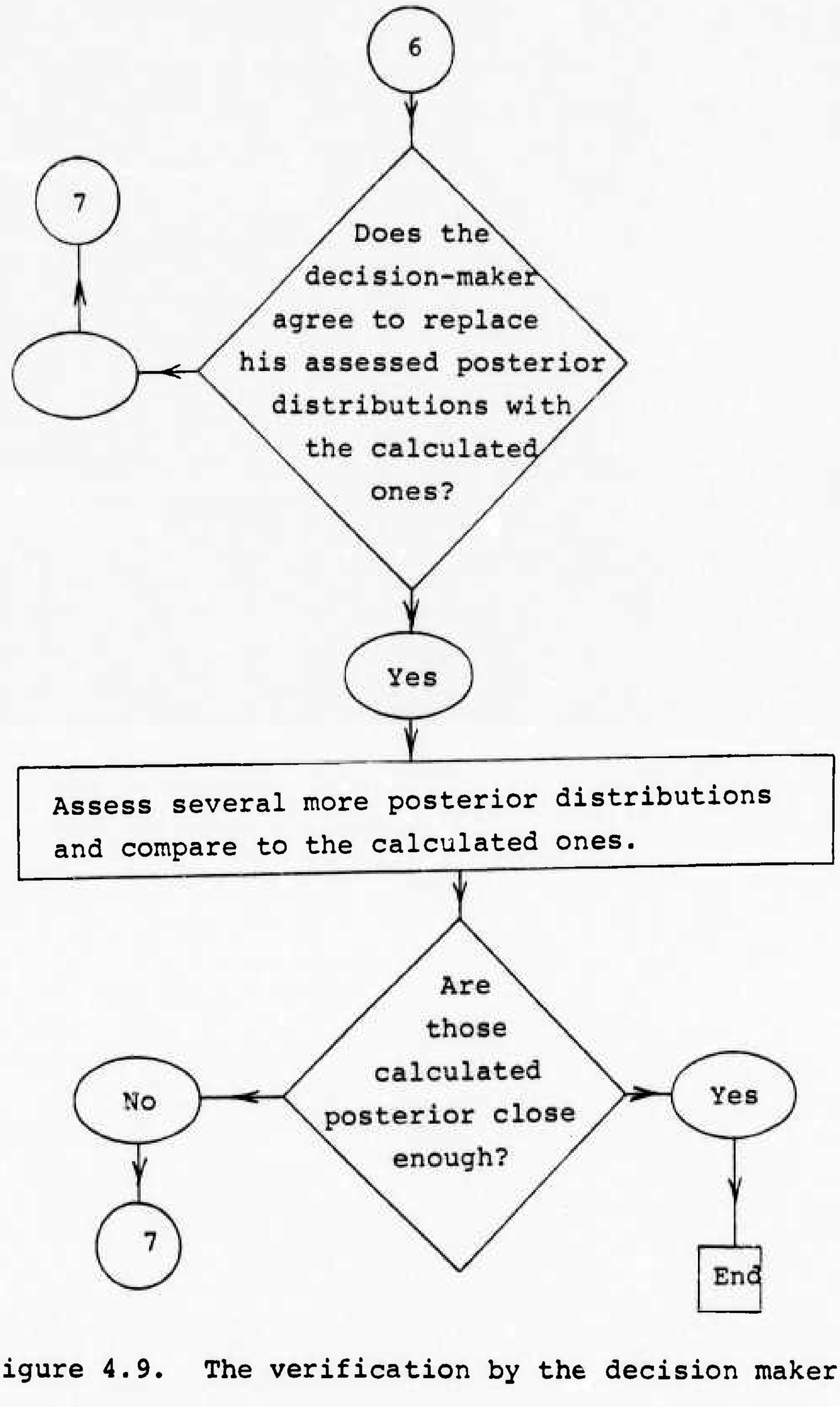


Figure 4.9. The verification by the decision maker

CHAPTER 5

THE ASSESSMENTS

5.1 Introduction

Our seven assessments are discussed in this chapter. The next section provides a general description of the assessments, including a specification of the drv's and a summary of the total errors for each. A detailed discussion of one of these assessments is presented in the last section.

5.2 An Overview of the Assessments

The seven drv's whose joint distributions we assessed are:

1. The yearly population of San Francisco County for the years 1961, 1962,..., 1971 as reported by the *California Statistical Abstract* of the California Department of Finance

2. The yearly average unemployment percentage in the United States for the years 1960, 1961,..., 1970 as reported by the *Economic Report of the President* (1971)

3. Annual earned income in 1976 dollars for the years Sept. 1976 - Sept. 1977, Sept. 1977 - Sept. 1978,..., Sept. 1985 - Sept. 1986

4. The yearly average price of U_3O_8 per pound for the years 1977, 1978,..., 1986 given that the current operating reactors are not shut down and plutonium recycle is allowed

5. Annual income in 1976 dollars from entrepreneurial activities for the years 1977, 1978,..., 1986

6. The price of one share of Golden Cycle stock at the close of business one week after the release of each Quarterly Report from July, 1976, through April, 1978

7. The yearly United States budget in 1976 dollars for the years 1977, 1978,..., 1986.

Table 5.1 presents some general information about each assessment. The first column lists the name of the drv. In the second column is the number M of posterior distributions that were assessed and used to determine $g_{0.5}$ and $g^*_{0.9}$. This number ranged between 12 and 18. The value of M_1 in the third column is the number of posterior distributions that were assessed but not used in the determination

TABLE 5.1

A Summary of the Assessments

Variable	N	M	M_1	ℓ
Population of San Francisco	11	15	5	3
Unemployment Percentage	11	18	5	2
Income	10	14	4	1
Price of U_3O_8	10	14	4	2
Entrepreneurial Income	10	17	4	2
Price of Golden Cycle Stock	8	12	4	2
U.S. Budget (10^9)	10	18	4	2

of the updating algorithms. These assessments were used to provide a check on the interpolation error. The order of conditional independence is listed in the last column. As stated in Chapter 4 this number varied between one and three for the drv's.

The errors for these assessments in terms of the average percent are presented in Table 5.2.

$$\text{error} = \sum_{j=1}^{M} \frac{|\delta - \hat{\delta}|}{\delta}$$

where δ was the assessed value

$\hat{\delta}$ was the approximate value.

The four categories that the error was computed for are the median, standard deviation, and two assessed fractiles (either 0.1/0.9 or 0.25/0.75). The top number for each assessment in each column is the error found in comparing the assessed posterior distributions that were used to determine the updating algorithms to their corresponding approximations using the updating algorithms. These numbers for the median and standard deviation reflect the curve-fitting error only. However, the errors for the outside fractiles are a combination of curve-fitting error and named distribution error. The numbers in parentheses are the associated errors for the four or five posterior distributions

TABLE 5.2

The Assessment Errors

Variable	Median	Standard Deviation	0.1/0.25 Fractile	0.75/0.9 Fractile
Population of San Francisco	0.005 (0.003)	0.109 (0.118)	0.019 (0.015)	0.006 (0.005)
Unemployment Percentage	0.023 (0.057)	0.179 (0.171)	0.031 (0.054)	0.027 (0.068)
Income	0.027 (0.031)	0.065 (0.132)	0.038 (0.044)	0.031 (0.038)
Price of U_3O_8	0.018 (0.024)	0.050 (0.142)	0.062 (0.033)	0.029 (0.071)
Entrepreneurial Income	0.048 (0.083)	0.072 (0.051)	0.087 (0.087)	0.044 (0.091)
Price of Golden Cycle Stock	0.015 (0.023)	0.057 (0.198)	0.053 (0.040)	0.036 (0.054)
U.S. Budget (10^9)	0.030 (0.072)	0.183 (0.309)	0.098 (0.172)	0.060 (0.012)

that were not used to determine the updating algorithm. These numbers in parentheses are our only check on interpolation error.

The errors for the posterior medians are very small in comparison to what the individuals considered to be their own assessment error. Also, the interpolation error in this case appears either insignificant or absent.

However, the curve-fitting error for the posterior standard deviations is larger than that for the medians in every case. In the second and seventh assessments it appears to border on the unacceptable. Some mitigating circumstances are that the individuals generally thought their assessment errors for the outside fractiles that determined the standard deviation were greater than the corresponding error for the medians. Also, they were more concerned that the general behavior of the standard deviation updating algorithm matched their beliefs than that these errors be minimized by using a lot of superfluous terms. In fact, once several of the individuals were happy with $g^*_{0.9}$, they were willing to change some of their assessments to match the approximation. Finally, as shown by the numbers in the last two columns, large errors in the posterior standard deviations do not necessarily mean large errors in the other fractiles of the distributions. For these reasons larger errors are acceptable in the standard deviation updating

algorithm and all of the individuals were willing to accept these errors. The numbers in parentheses indicate that in some of these cases the interpolation error may be large. However, these numbers do not reflect interpolation error alone and are influenced by all of the above factors.

The errors in the outside fractiles are acceptable considering assessment error. This indicates that the error introduced into a decision analysis by using these approximated posterior distributions is within the acceptable region. In fact, all of the individuals said they would be willing to use these approximations.

5.3 The Assessment of the Price of U_3O_8

A detailed discussion of the assessment of a joint probability distribution on the following drv is provided in this section:

> The yearly average price of U_3O_8 per pound for the years 1977, 1978,..., 1986 given that the current operating nuclear reactors are not shut down and plutonium recycle is allowed.

U_3O_8 is a uranium oxide ore that is mined and purified for use as a fuel for nuclear power plants. The individual who played the decision maker or expert role for these assessments is a nuclear engineer. While this drv was not part of a decision analysis, it is clearly the type of variable for which a joint distribution might be required.

Note that the yearly price is conditioned on two events: (1) the current operating nuclear reactors are not shut down and (2) plutonium recycle is allowed. This is the type of modeling that was referred to in Chapters 3 and 4. The expert initially felt that the drv should be conditioned on the first event above in order to keep the joint distribution free of wild types of behavior. Then during the final assessment phase his assessed posterior distributions were exhibiting certain erratic forms of behavior that were traced to the resolution of the plutonium recycle question around the year 1980. For this reason the drv was conditioned on this event also. In an actual decision analysis a sensitivity analysis of the decision to these newly defined drv's would be performed to determine whether two, one, or no joint probability distribution(s) should be assessed.

Since the expert was very familiar with the theories of decision analysis and probability, the initial steps of the preliminary assessment procedure were dispensed with quickly. The prior distributions were encoded for the years 1977

through 1986 and are represented by the 0.1, 0.5, and 0.9 fractiles in Figure 5.1. As can be seen in this figure, the price of U_3O_8 is generally expected to rise. The prior distributions are not exactly symmetric but are close. Next, the order of conditional independence was ascertained to equal two. During this assessment process the expert was not asked to draw the graphs that are part of the final stages of the preliminary assessment phase. However, we did discuss the essence of these graphs so that they can be reconstructed here in general terms. The posterior median of $x(i)$ was thought to be an increasing function of $x(i-1)$ as shown by the solid line in Figure 5.2. The expert felt that the prior fractile paths were very likely paths for the drv to follow over time and that the posterior median of $x(i)$ was a function of the trend of the two previous prior fractiles. That is, the posterior median of $x(i)$ is also an increasing function of the difference between the revealed prior fractiles of $x(i-1)$ and $x(i-2)$. This is indicated by the three dashed lines in Figure 5.2 which represent different values of $x(i-2)$. This same relationship held for the posterior variance.

The first step of the preliminary analysis phase is to approximate the assessed prior distributions by the appropriate members of a named distribution. The normal distribution provided a reasonably accurate approximation of the prior distributions, and Table 5.3 shows the median (mean)

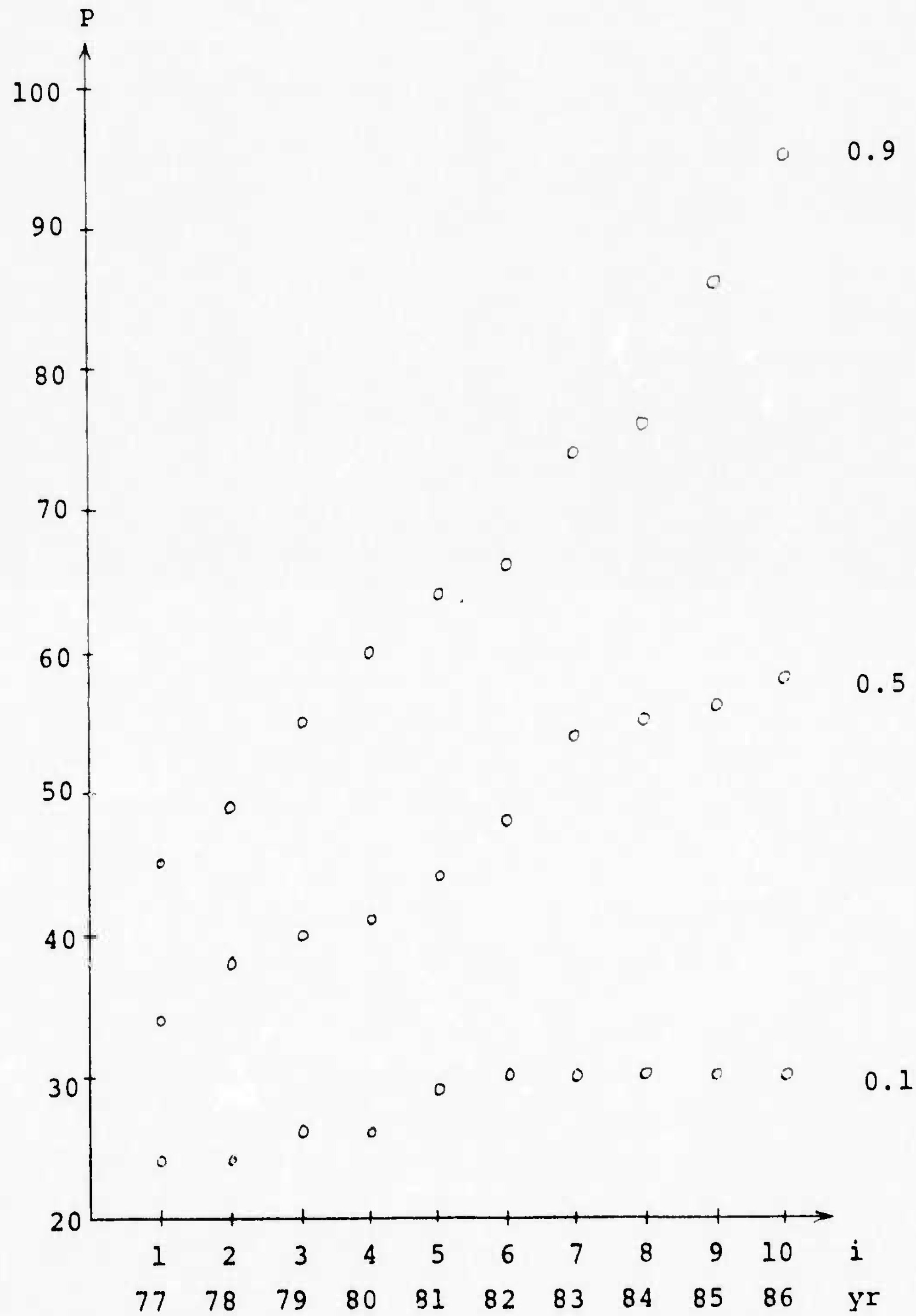


Figure 5.1. The prior fractiles on the price of U_3O_8

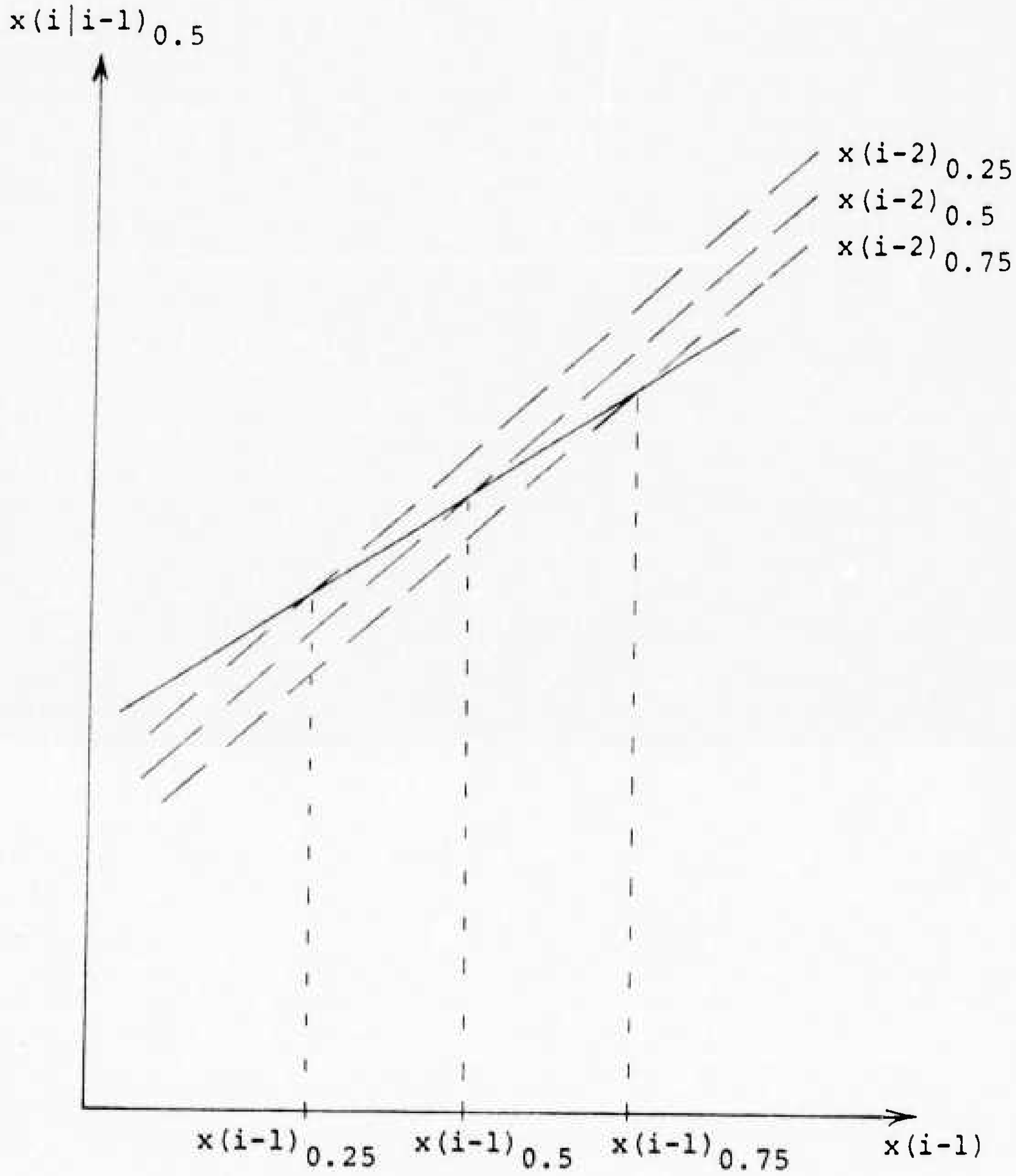


Figure 5.2. The expert's preliminary graphs of the posterior median versus $x(i-1)$ and $x(i-2)$

TABLE 5.3

The Approximate Prior Distributions

Year	Mean	Standard Deviation
1977	34.	8.2
1978	38.	9.8
1979	40.	11.
1980	41.	13.
1981	44.	14.
1982	48.	14.
1983	54.	17.
1984	55.	18.
1985	56.	22.
1986	58.	25.

and standard deviation of the approximate prior distributions. Next preliminary forms of the $g_{0.5}$ and $g^{*}_{0.9}$ functions were postulated. The simplest linear form is:

$$a_0+a_1 f(i-1)+a_2(f(i-1)-f(i-2)). \qquad (5.1)$$

However, other nonlinear forms such as quadratic or logarithmic forms could be used if needed as long as the result is an increasing function of $f(i-1)$ and $(f(i-1)-f(i-2))$. Also, the last term of equation 5.1 can be multiplied by the following expression which was discussed in Chapter 4.3:

$$\begin{array}{ll} 0.5-a_0-a_1 f(i-1), & f(i-1)>f(i-2) \\ a_0+a_1 f(i-1)+0.5, & f(i-1)<f(i-2). \end{array}$$

Finally, the analyst must choose which posterior distributions to assess in the next phase. The posterior distributions encoded in this assessment are listed in Table 5.4.

The posterior assessment phase proceeded smoothly once the yearly average price per pound of U_3O_8 was conditioned on the two events mentioned above. Since the explanation of the graphical technique for defining the functional forms and parameters of $g_{0.5}$ and $g^{*}_{0.9}$ requires only eight of these assessments, Table 5.5 presents the results of the first eight posterior distributions in Table 5.4.

TABLE 5.4

The Given Information for the

Assessed Posterior Distributions

Year	i	x(i-2)	x(i-1)
1978	2	-	34.
1978	2	-	36.
1978	2	-	40.
1978	2	-	30.
1979	3	35.	38.
1979	3	35.	35.
1979	3	30.	35.
1979	3	35.	32.
1980	4	38.	40.
1980	4	38.	44.
1980	4	41.	44.
1980	4	44.	42.
1984	8	48.	50.
1984	8	54.	54.

TABLE 5.5

8 Assesed Posterior Distributions

i	$x(i-2)$	$x(i-1)$	$x(i\|i-1)_{0.1}$	$x(i\|i-1)_{0.5}$	$x(i\|i-1)_{0.9}$	$\mu(i\|i-1)$	$\sigma(i\|i-1)$	$g_{0.5}$	$g^*_{0.9}$
2	-	34.	28.	36.	44.	36.	6.2	-.08	-.11
2	-	36.	28.	40.	46.	40.	7.0	.08	-.08
2	-	40.	32.	45.	52.	45.	7.8	.26	-.05
2	-	30.	25.	30.	40.	30.	5.9	-.29	-.12
3	35.	38.	30.	41.	55.	41.	9.8	.03	-.04
3.	35.	35.	29.	36.	48.	36.	7.4	-.14	-.10
3	30.	35.	26.	37.	52.	37.	11.	-.10	-.02
3	35.	32.	26.	31.	40.	31.	5.5	-.29	-.17

All of the assessed posterior distributions were approximated by normal distributions in the first step of the final analysis phase. These normal distributions are also tabulated in Table 5.5 for the first eight posterior distributions. Next the values of $g_{0.5}$ and $g^*_{0.9}$ are calculated for each of these distributions using equations 3.4 and 3.11. These values are also shown in Table 5.5.

The next step in this phase of the assessment procedure is to use the assessed posterior distributions to completely define the forms and parameters of $g_{0.5}$ and $g^*_{0.9}$. As was mentioned in Chapter 4.5, this can be done by using least squares to find the parameters for several different forms of these models. However, we have used a graphical procedure that provides both the intuitive and computational information needed by the analyst for choosing the forms and parameters of the two models. First $g_{0.5}$ is plotted versus $f(1)$ for the four assessed posterior distributions in 1978. This graph is shown in Figure 5.3. The four distributions are represented by dots. The straight line

$$-0.62+1.12\, f(1)$$

provides a close approximation. This approximation would not be improved much by a nonlinear model so none was investigated. Next we plot $g_{0.5}$ versus $f(2)$ for the three posterior distributions on $x(3)$ which have a given

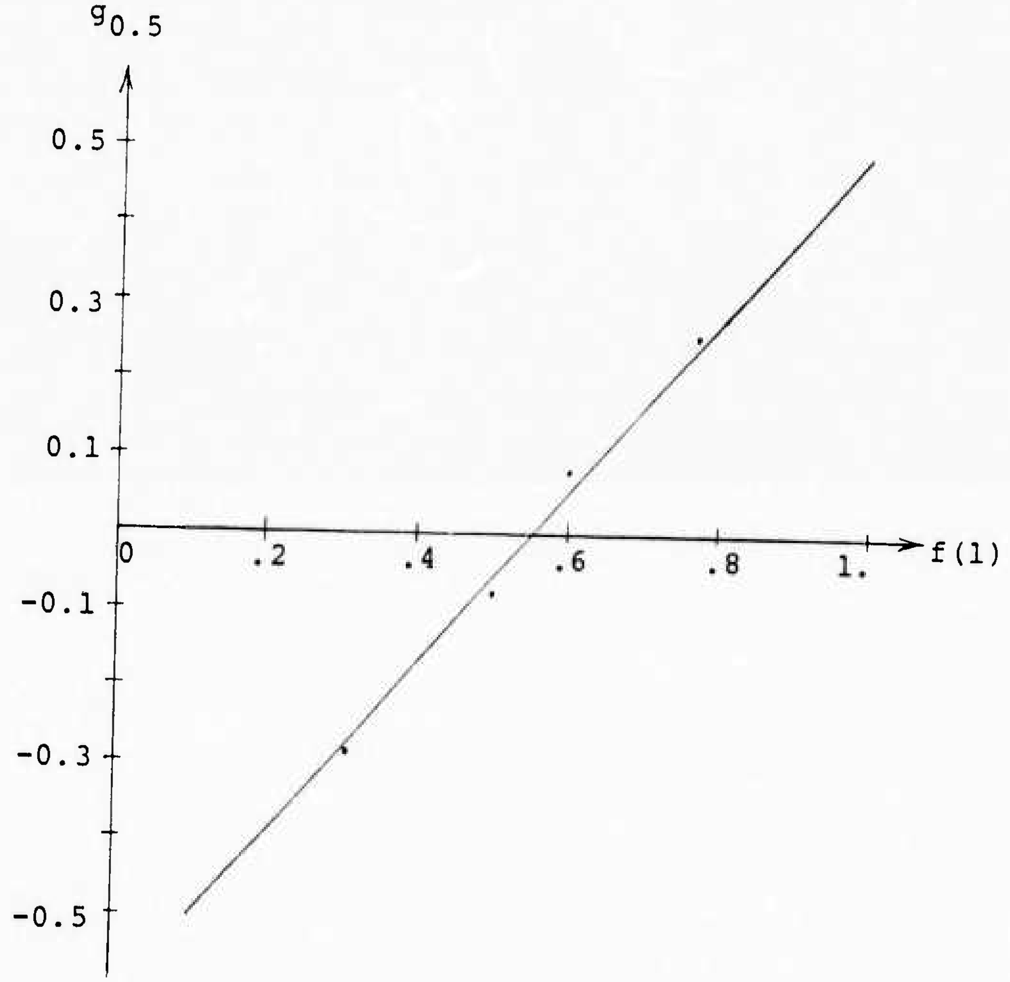


Figure 5.3 $g_{0.5}$ versus f(1) for i=2

value of $x(1)$ equal to 35. The dots in Figure 5.4 represent these three distributions. The prior fractile of $x(1)$ corresponding to 35 is 0.55. These three points are very accurately approximated by the solid straight line shown. In order to be consistent with the statements of the preliminary assessment phase, this solid line must intersect the line representing $g_{0.5}$ when $f(1) = f(2)$ at $f(2)$ equal to 0.55. The dotted line in Figure 5.4 represents $g_{0.5}$ when $f(1) = f(2)$ and has the same slope although different intercept as the line in Figure 5.3. The last posterior distribution on $x(3)$ is represented by the circle in Figure 5.4. It corresponds to $f(1) = 0.31$ and $f(2) = 0.38$, and should lie slightly above the dotted line rather than slightly below. However, this is still a reasonable approximation. These assessments indicate that the linear form of $g_{0.5}$ is very adequate. The final form of $g_{0.5}$ for every period is as follows:

$$
\begin{aligned}
g_{0.5} &= a_1(i)+1.12f(i-1)+.29(f(i-1)-f(i-2)) \\
a_1(i) &= \begin{cases} -.62 & i=2 \\ -.52 & i=3 \\ -.59 & i=4 \\ -.56 & i>4 \end{cases} \qquad (5.2) \\
f(0) &= f(1) \text{ when } i=2.
\end{aligned}
$$

Note that only the intercept is a function of i. The value of $a_1(i)$ for i greater than 4 was ascertained by the

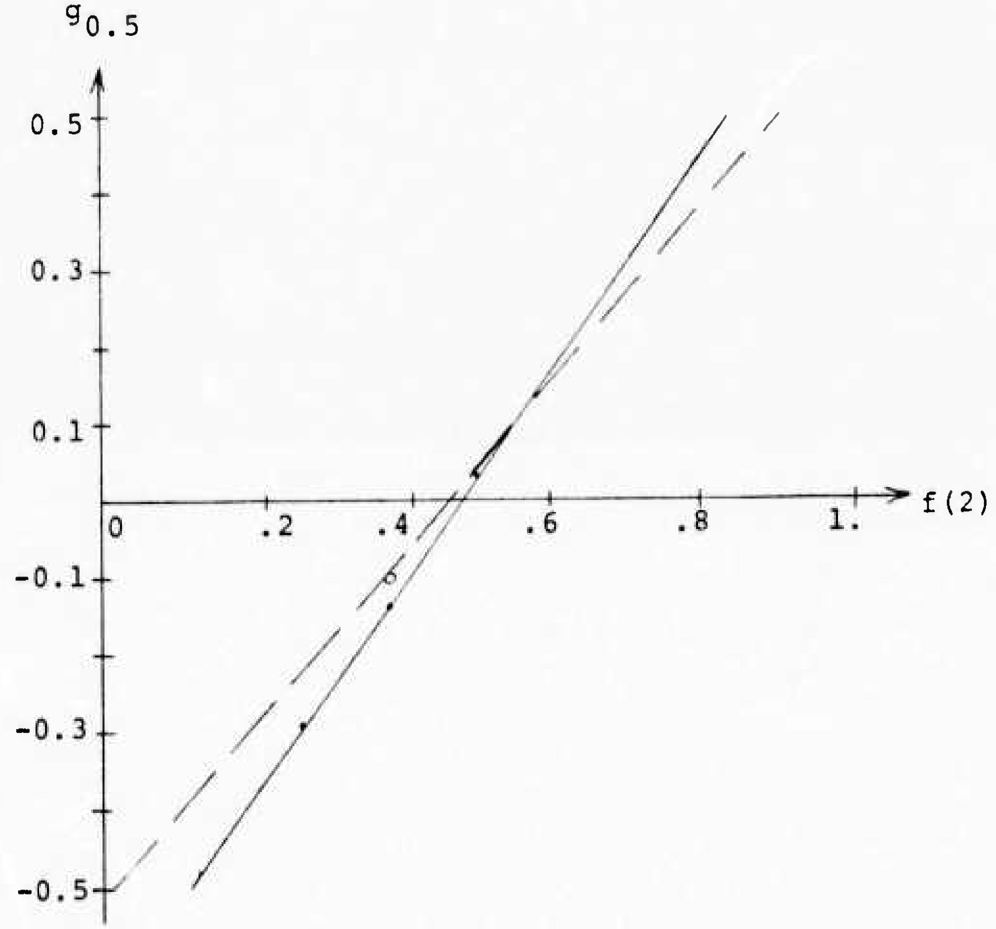


Figure 5.4. $g_{0.5}$ versus f(2) for i=3.

last two assessments and information gathered in the preliminary assessment phase. The expert stated that the posterior median in periods five through ten, given that the previous path was the prior medians, was equal to the corresponding prior median. Finally, it should be noted that there are values of $f(i-2)$ and $f(i-1)$ for which $g_{0.5}$, as stated above, violates its constraints. Since the linear form is so accurate for nonextreme values of $f(i-1)$ and $f(i-2)$, we decided to add a condition that would modify the value of $g_{0.5}$ only when it was near or in actual violation of a constraint. For instance we used the following scheme:

$$\begin{aligned}
&\text{if } g_{0.5} > 0.4,\ g_{0.5} = (g_{0.5} - 0.4)(\max g_{0.5} - 0.4) \qquad (5.3)\\
&\text{if } g_{0.5} < -0.4,\ g_{0.5} = -(-g_{0.5} + 0.4)(\min g_{0.5} + 0.4).
\end{aligned}$$

This insures that the constraints are met.

The procedure for determining $g^*_{0.9}$ is exactly the same. Figure 5.5 shows $g^*_{0.9}$ plotted versus $f(1)$ for the four posterior distributions on $x(2)$. Again a straight line seems to be the best approximation. The straight line depicted is

$$-0.175 + 0.15\ f(1).$$

The three dots in Figure 5.6 represent the posterior distributions on $x(3)$ when $x(1)$ equals 35 (or $f(1) = 0.55$).

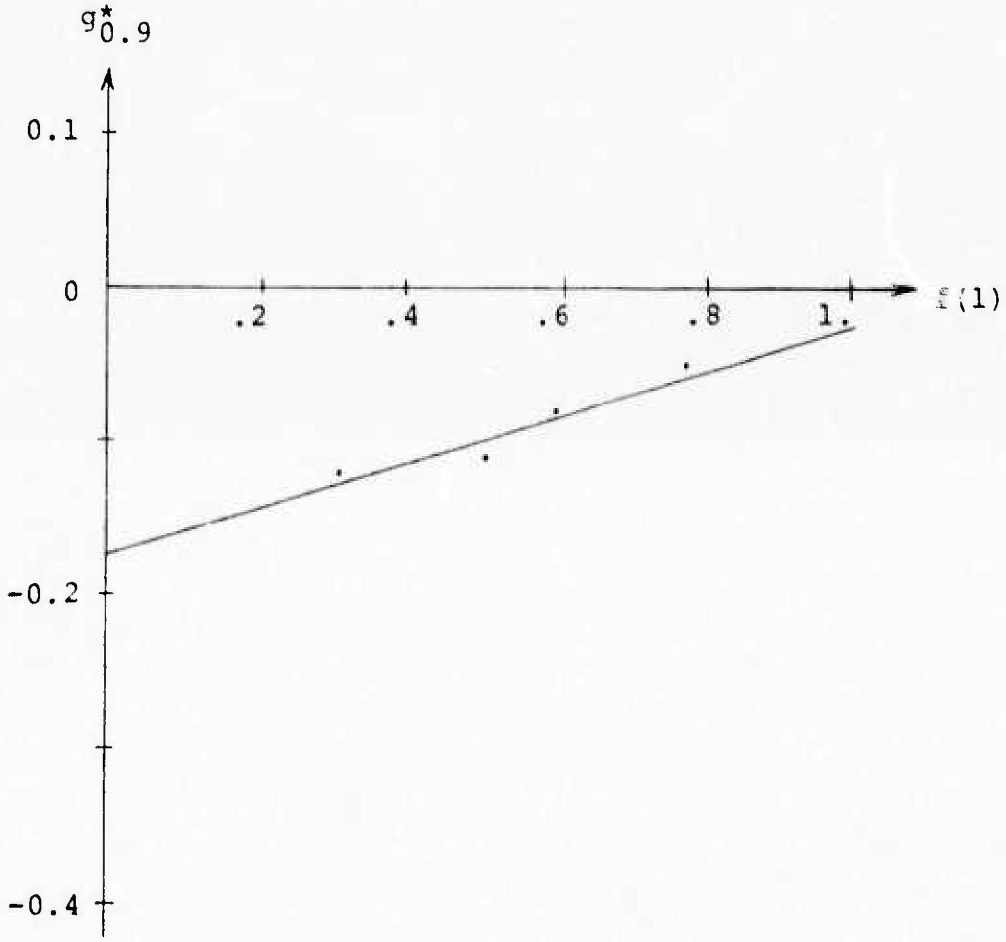


Figure 5.5 $g^*_{0.9}$ versus f(1) for i=2

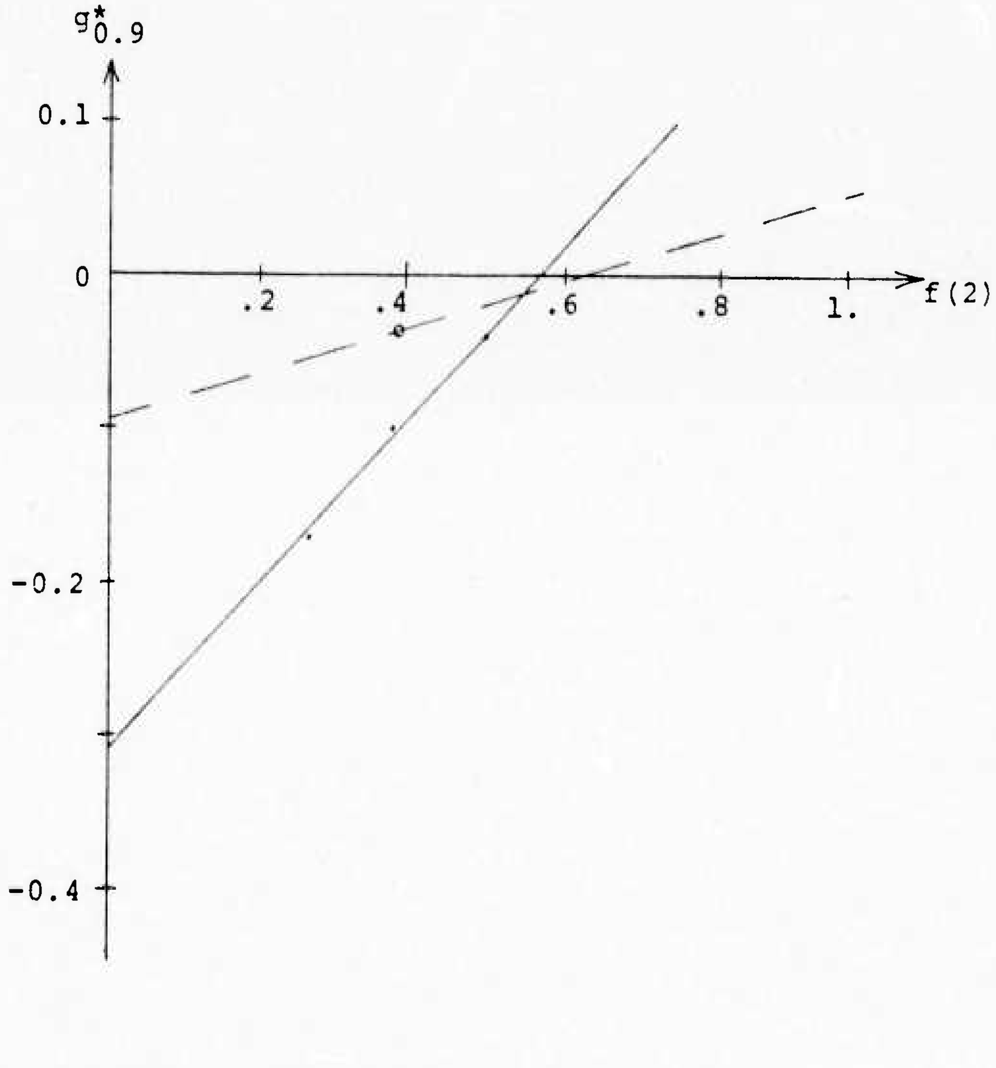


Figure 5.6 $g^{*}_{0.9}$ versus f(2) for i=3

They also are closely approximated by a straight line. Again the dashed line in this figure intersects the solid line at $f(2)$ equal to 0.55 and has the same slope as the line in Figure 5.5. The circle in Figure 5.6 represents the last posterior distribution on $x(3)$, with $f(1) = 0.31$ and $f(2) = 0.38$. In this case it lies directly on the dashed line rather than slightly above it. The final form of $g^*_{0.9}$ is:

$$g^*_{0.9} = a_2(i)+.15f(i-1)+.35(f(i-1)-f(i-2))$$

$$a_2(i) = \begin{cases} -.175 & i=2 \\ -.095 & i=3 \\ -.235 & i>3 \end{cases} \qquad (5.4)$$

$$f(0) = f(1) \text{ when } i=2.$$

Again the intercept is the only parameter that is a function of the period. The constraints are much more likely to be violated for $g^*_{0.9}$. But since the linear form is again very accurate, the same strategy as was used for $g^*_{0.9}$ was used to meet the constraints.

The next step is to compute the approximated assessed posterior distributions and compare them with the assessments. Table 5.6 presents this comparison. The last four distributions are additional assessments that were done at the same time as the other posterior distributions but which were not used to determine the forms or parameters of $g_{0.5}$

TABLE 5.6

Comparison of Assessed and
Approximate Posterior Distributions

Assessed Median	Approximate Median	Assessed Standard Deviation	Approximate Standard Deviation
36	37	6.2	6.4
40	39	7.0	6.8
45	44	7.8	7.5
30	31	5.9	5.6
41	41	9.8	9.6
36	36	7.4	7.5
37	38	10.6	10.6
31	31	5.5	5.9
40	40	6.2	6.7
47	47	9.0	9.1
46	46	7.4	7.6
39	41	6.6	5.3
51	49	7.4	7.2
54	53	7.4	6.7
44	44	8.6	10.0
35	34	6.6	5.4
56	55	8.6	9.0
60	57	10.2	8.4

and $g^*_{0.9}$. The approximate and assessed medians are very close, the maximum error being five percent. The approximate and assessed standard deviations are also close with a few significant errors. The expert stated that he felt these comparisons and the general behavior of $g_{0.5}$ and $g^*_{0.9}$ captured his uncertainty with enough accuracy that he would use this joint probability distribution in a decision analysis if it were needed and he were the decision maker.

CHAPTER 6

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

6.1 Conclusions

We began by examining the problem of assessing uncertainties on dynamic random variables. These variables are one of the basic building blocks in models of decisions involving dynamic situations. The previous work in this area is scarce, and this research presents a method for dealing with the complexities of this problem.

Multivariate named distributions were demonstrated to be inadequate approximations of decision makers' joint distributions in general. This inadequacy is a result of the rigid updating implications of these named distributions. A decision maker would have to have very special beliefs about how his prior distributions are updated to posterior distributions to be willing to use either the normal, lognormal, or Student distributions.

A method that approximates the decision maker's uncertainty by modeling his updating beliefs was presented in Chapter 3. This method was designed to approximate accurately almost any probabilistic beliefs a decision-maker

might have about a drv. It was also designed so that the assessment procedure could be as short as possible and use state-of-the-art encoding techniques. Finally, it has a robust mathematical framework so that only a reasonable number of assessments are required to specify fully the approximation of the decision maker's beliefs. The assessments we have done verify that this method does indeed provide useful and accurate approximations of joint probability distributions.

Chapter 4 describes the assessment procedure that was refined in the process of doing our assessments. It has proven to be helpful to the decision maker in thinking about his beliefs. It also provides a systematic process for the analyst to determine the approximation of these beliefs and gather the information needed to quantify them. The example in Chapter 5 illustrates this procedure in detail.

6.2 Suggestions for Further Research

This research is a first step into the realm of quantifying uncertainty on dynamic random variables. It is hoped that this research will become a point of departure for future research in this area. First it will be an aid to understanding the problem, but, more importantly, a tool for solving the problem so that experience can be gained in dealing with complex stochastic systems. By using the

method developed in this research, the analyst will gain insight into the forces which drive different types of probabilistic beliefs. This experience should generate a better understanding of the proprieties of various types of stochastic models for generating the decision maker's probabilistic beliefs. This is the more direct but difficult solution to this problem. Currently, only linear stochastic dynamic models of a dynamic random variable(s) are used with any frequency and understood completely. However, as stated in Chapter 1 they are the equivalent of a multivariate normal distribution which few decision makers would accept.

To complement this experience, future researchers should explore the recent research on nonlinear filtering theory. This research does not provide the answers needed in decision analysis because of the difference in the problems with which each are concerned. The control engineer typically models systems in-the-small and has to be able to handle a great deal of data in a short period of time. Conversely, the decision analyst models systems in-the-large and has to make the best use of the little data he can find. However, it is possible that a thorough understanding of nonlinear filtering theory and more experience with dynamic random variables in future decision analyses will provide valuable insights to this problem.

APPENDIX A

NOTATION

$a_{(\cdot)}$ - a constant

$c(i)$ - the consumption in period i

C - a normalizing constant for the Student distribution

f - the fractile of a probability distribution, $0\leq f\leq 1$

$f(j)$ - the prior fractile of the revealed value of the drv in period j

$f_e(x|\lambda)$ - the density function of the exponential distribution

$f_{ln}^{(N)}(\underline{x}|\underline{\mu},\underline{\Sigma})$ - the multivariate lognormal density function of dimension N. If $N=1$, the superscript is omitted.

$f_n^{(N)}(\underline{x}|\underline{\mu},\underline{\Sigma})$ - the multivariate normal density function of dimension N

$f_S^{(N)}(\underline{x}|n,\underline{\mu},\underline{T})$ - the multivariate Student density function of dimension N

$F_{(\cdot)}(x|\cdot,\cdot)$ - the $(\cdot)$ cumulative function

g_f - the delta in fractile units that satisfies

$x(i|\xi)_{g_f+f} = x(i|i-1)_f$

g_f^* - the distance in fractile units from the prior f fractile to the posterior cumulative

$h_1(\cdot)$ - a function

$h_2(\cdot,\cdot)$ - a function

i - a discrete time indicator

j - a discrete time indicator

k - a discrete time indicator

ℓ - the order of conditional independence

n - the degree of freedom parameter of the Student distribution

N - the number of discrete time periods

$\underline{P}$ - the precision matrix of the multivariate normal distribution

q - a discrete time indicator

t - continuous time

$\underline{T}$ - the precision matrix of the multivariate Student distribution

$\underline{W}$ - the covariance matrix of the multivariate Student distribution

$\underline{x}$ - the N-dimensional vector of values of the drv

$x(i)$ - the value of the drv in period i

y - the standardized normal variate, $f_n(y|0,1)$

$\underline{z}$ - an N-dimensional variable

$z(i)$ - the value of z in period i

α - a parameter of the Weibull distribution

β - a parameter of the Weibull distribution

$\Gamma(\cdot)$ - the gamma function

$\underline{\theta}$ - the N-dimensional vector of values of an unobservable drv

$\theta(i)$ - the value of $\underline{\theta}$ in period i

λ - a parameter of the exponential distribution

$\underline{\mu}$ - the N-dimensional vector of the normal, lognormal, Student distributions

π - the constant 3.1415...

ρ_{ij} - the correlation coefficient between the i and j values for the multivariate normal distribution

$\underline{\Sigma}$ - the covariance matrix of the normal and lognormal distributions

τ - a length of continuous time

$(\cdot)'$ - the transpose of the vector or matrix $(\cdot)$

The following is a description of the inferential notation:

$\mathcal{S}$ - a state of information

ξ - the prior state of information

$\{x|\mathcal{S}\}$ - the density function on x given state of information $\mathcal{S}$

$\langle x|\mathcal{S}\rangle$ - the expected value of x given state of information $\mathcal{S}$

${}^{v}\langle x|\mathcal{S}\rangle$ - the variance of x given state of information $\mathcal{S}$

$(\cdot)_f$ - the f fractile of $(\cdot)$

$(\cdot)_\phi$ - a dummy variable

$x(i|\xi)_f$ - the prior f fractile of $x(i)$

$x(i|i-1)_f$ - the posterior f fractile of $x(i)$ given the values of $x(1),\ldots,x(i-1)$

$\overset{v}{x}(i|\xi)$ - the prior variance of $x(i)$

$\overset{v}{x}(i|i-1)$ - the posterior variance of $x(i)$ given the values of $x(1),\ldots,x(i-1)$

$s_{x(i|\xi)}$ - the prior standard deviation of $x(i)$

$s_{x(i|i-1)}$ - the posterior standard deviation of $x(i)$ given the values of $x(1),\ldots,x(i-1)$

APPENDIX B

PROOF OF EQUATION 2.7

This appendix gives a proof of the following proposition: Let N random variables, $x(1), x(2), \ldots, x(N)$, be described by any joint probability distribution $\{x(1), x(2), \ldots, x(N) | \xi\}$. Define the posterior mean of any $x(i)$ given a subset j $(1<j<N-1)$ of the remaining $N-1$ random variables.

$$\langle x(i) | x(1), x(2), \ldots, x(j), \xi \rangle \quad (i=2,\ldots,N), (i>j).$$

Note that we can denote this subset as $x(1), x(2), \ldots, x(j)$ without any loss of generality. Then if this posterior mean is a linear combination of $x(1), x(2), \ldots, x(j)$:

$$\langle x(i) | x(1), x(2), \ldots, x(j), \xi \rangle (x(1)_\phi, x(2)_\phi, \ldots, x(j)_\phi) = a_0 + a_1 x(1)_\phi + a_2 x(2)_\phi + \ldots + a_j x(j)_\phi,$$

where $a_0, a_1, \ldots, a_j$ are constants, the prior mean of $x(i)$ is equal to this linear combination, evaluated at the prior means of $x(1), x(2), \ldots, x(j)$.

$$\langle x(i) | \xi \rangle = a_0 + a_1 \langle x(1) | \xi \rangle + a_2 \langle x(2) | \xi \rangle + \ldots + a_j \langle x(j) | \xi \rangle.$$

The proof of this is a direct integration of the appropriate multiple integral. The prior mean is defined to be

$$\langle x(i)|\xi\rangle = \int_{x(1)}\int_{x(2)}\cdots\int_{x(N)} x(i)\{x(1),x(2),\ldots,x(N)|\xi\}.$$

This can easily be reduced to

$$\begin{aligned}
\langle x(i)|\xi\rangle &= \int_{x(1)}\int_{x(2)}\cdots\int_{x(j)}\int_{x(i)} x(i)\{x(1),x(2),\ldots,\\
&\quad x(j),x(i)|\xi\}\\
&= \int_{x(1)}\int_{x(2)}\cdots\int_{x(j)}\int_{x(i)} x(i)\{x(i)|x(j),\ldots,\\
&\quad x(1),\xi\}.\quad \{x(1),\ldots,x(j)|\xi\}\\
\langle x(i)|\xi\rangle &= \int_{x(1)}\int_{x(2)}\cdots\int_{x(j)} \langle x(i)|x(j),\ldots,x(1),\xi\rangle.\\
&\quad \{x(1),\ldots,x(j)|\xi\}\\
&= \int_{x(1)}\int_{x(2)}\cdots\int_{x(j)} (a_0+a_1x(1)+\ldots+a_jx(j)).\\
&\quad \{x(1),\ldots,x(j)|\xi\}\\
&= a_0+a_1\int_{x(1)} x(1)\{x(1)|\iota\}+\ldots+a_j\int_{x(j)} x(j)\{x(j)|\xi\}\\
&= a_0+a_1\langle x(1)|\xi\rangle+\ldots+a_j\langle x(j)|\xi\rangle.
\end{aligned}$$

APPENDIX C

THE POSTERIOR MULTIVARIATE STUDENT DENSITY FUNCTION

We are given the fact that

$$\{\underline{x}|\xi\} = f_S^{(N)}(\underline{x}|n,\underline{\mu},\underline{T})$$

$$= C(1+\frac{1}{n}(\underline{x}-\underline{\mu})'\underline{T}(\underline{x}-\underline{\mu}))$$

where

$$C = \frac{\Gamma\left(\frac{n+N}{2}\right)\left|\underline{T}\right|^{\frac{1}{2}}}{\Gamma\left(\frac{n}{2}\right)(n\pi)^{\frac{N}{2}}} .$$

We wish to calculate $\{\underline{x}_2|\underline{x}_1,\xi\}$ where $\underline{x}_1$ has dimension q and $\underline{x}_2$ has dimension $N-q$. We know by the definition of conditional probability that

$$\{\underline{x}_2|\underline{x}_1,\xi\} = \frac{\{\underline{x}_1,\underline{x}_2|\xi\}}{\{\underline{x}_1|\xi\}} = \frac{\{\underline{x}|\xi\}}{\{\underline{x}_1|\xi\}} .$$

The probability distribution $\{\underline{x}_1|\xi\}$ equals

$$\{\underline{x}_1|\xi\} = f_S^{(q)}(\underline{x}_1|n,\underline{\mu}_1,\underline{T}_{11}-\underline{T}_{12}\underline{T}_{22}^{-1}\underline{T}_{21})$$

$$= f_S^{(q)}(\underline{x}_1|n,\underline{\mu}_1,\underline{W}_{11}^{-1}) .$$

Therefore

$$\{\underline{x}_2 | \underline{x}_1, \xi\} = C' \frac{\left[1+\frac{1}{n}(\underline{x}-\underline{\mu})'\underline{T}(\underline{x}-\underline{\mu})\right]^{-\left(\frac{n+N}{2}\right)}}{\left[1+\frac{1}{n}(\underline{x}_1-\underline{\mu}_1)'\underline{W}_{11}^{-1}(\underline{x}_1-\underline{\mu}_1)\right]^{-\left(\frac{n+q}{2}\right)}} \tag{C.1}$$

where

$$C' = \frac{\dfrac{\Gamma\left(\frac{n+N}{2}\right) \left|\underline{T}\right|^{\frac{1}{2}}}{\Gamma\left(\frac{n}{2}\right) (n\pi)^{\frac{N}{2}}}}{\dfrac{\Gamma\left(\frac{n+q}{2}\right) \left|\underline{W}_{11}^{-1}\right|^{\frac{1}{2}}}{\Gamma\left(\frac{n}{2}\right) (n\pi)^{\frac{q}{2}}}} = \frac{\Gamma\left(\frac{n+N}{2}\right) \left|\underline{T}\right|^{\frac{1}{2}} \left|\underline{W}_{11}\right|^{-\frac{1}{2}}}{\Gamma\left(\frac{n+q}{2}\right) (n\pi)^{\frac{N-q}{2}}} .$$

Note that $(1+\frac{1}{n}(\underline{x} - \underline{\mu})'\underline{T}(\underline{x} - \underline{\mu}))$ can be rewritten as

$$(1+\frac{1}{n}(\underline{x}_1-\underline{\mu}_1)'\underline{W}_{11}^{-1}(\underline{x}_1-\underline{\mu}_1)+$$

$$\frac{1}{n}((\underline{x}_2-\underline{\mu}_2)+\underline{T}_{22}^{-1}\underline{T}_{21}(\underline{x}_1-\underline{\mu}_1))'\underline{T}_{22}((\underline{x}_2-\underline{\mu}_2)+\underline{T}_{22}^{-1}\underline{T}_{21}(\underline{x}_1-\underline{\mu}_1))).$$

Substituting this into equation C.1 and multiplying by

$$1 = \frac{\left[1+\frac{1}{n}(\underline{x}_1-\underline{\mu}_1)'\underline{W}_{11}^{-1}(\underline{x}_1-\underline{\mu}_1)\right]^{\frac{-(N-q)}{2}}}{\left[1+\frac{1}{n}(\underline{x}_1-\underline{\mu}_1)'\underline{W}_{11}^{-1}(\underline{x}_1-\underline{\mu}_1)\right]^{\frac{-(N-q)}{2}}}$$

we get

$$\{\underline{x}_2 | \underline{x}_1, \xi\} = C'' \left[\frac{1+\frac{1}{n}(\underline{x}_1-\underline{\mu}_1)'\underline{W}_{11}^{-1}(\underline{x}_1-\underline{\mu}_1)}{1+\frac{1}{n}(\underline{x}_1-\underline{\mu}_1)'\underline{W}_{11}^{-1}(\underline{x}_1-\underline{\mu}_1)} + \right.$$

$$\left. \frac{\frac{1}{n}((\underline{x}_2-\underline{\mu}_2)+\underline{T}_{22}^{-1}\underline{T}_{21}(\underline{x}_1-\underline{\mu}_1))'\underline{T}_{22}((\underline{x}_2-\underline{\mu}_2)+\underline{T}_{22}^{-1}\underline{T}_{21}(\underline{x}_1-\underline{\mu}_1))}{1+\frac{1}{n}(\underline{x}_1-\underline{\mu}_1)'\underline{W}_{11}^{-1}(\underline{x}_1-\underline{\mu}_1)} \right]^{-\frac{(n+N)}{2}}$$

where $C'' = C'(1+\frac{1}{n}(\underline{x}_1-\underline{\mu}_1)'\underline{W}_{11}^{-1}(\underline{x}_1-\underline{\mu}_1))^{-\frac{(N-q)}{2}}$. Thus

$$\{\underline{x}_2 | \underline{x}_1, \xi\} = C'' \left[1+ (\underline{x}_2-\underline{\mu}_{2|1})' \frac{\underline{T}_{22}}{n+(\underline{x}_1-\underline{\mu}_1)'\underline{W}_{11}^{-1}(\underline{x}_1-\underline{\mu}_1)} (\underline{x}_2-\underline{\mu}_{2|1}) \right]^{-\frac{(n+N)}{2}}$$

$$= C'' \left[1+\frac{1}{n+q} (\underline{x}_2-\underline{\mu}_{2|1})' \frac{(n+q)\underline{T}_{22}}{n+(\underline{x}_1-\underline{\mu}_1)'\underline{W}_{11}^{-1}(\underline{x}_1-\underline{\mu}_1)} (\underline{x}_2-\underline{\mu}_{2|1}) \right]^{-\frac{(n+q+N-q)}{2}}$$

$$= f_S^{(N-q)}\left(\underline{x}_2 | n+q, \underline{\mu}_{2|1}, \frac{(n+q)\underline{T}_{22}}{n+(\underline{x}_1-\underline{\mu}_1)'\underline{W}_{11}^{-1}(\underline{x}_1-\underline{\mu}_1)} \right)$$

where $\underline{\mu}_{2|1} = \underline{\mu}_2 - \underline{T}_{22}^{-1}\underline{T}_{21}(\underline{x}_1-\underline{\mu}_1)$.

APPENDIX D

A DESCRIPTION OF A GENERAL METHOD FOR UPDATING

As stated in Chapter 3, the function g is the vertical difference from the posterior cumulative to the prior at the f fractile of the posterior distribution. (See Figure D.1.) Therefore g is also defined by the equation

$$x(i|i-1)_f = x(i|\xi)_{f+g} \tag{D.1}$$

or the prior $(f+g)$ fractile of $x(i)$ equals the posterior f fractile of $x(i)$. The value of g depends on f as well as the posterior information.

Some of the outstanding properties of g are best illustrated by examining it for the two-dimensional multivariate normal distribution:

$$\begin{aligned}\{\underline{x}|\xi\} &= \{x(1),x(2)|\xi\} \\ &= \frac{1}{2\pi}\left|\Sigma\right|^{\frac{1}{2}} \exp\left[-\frac{1}{2}(\underline{x}-\underline{\mu})'\underline{\Sigma}^{-1}(\underline{x}-\underline{\mu})\right]\end{aligned} \tag{D.2}$$

where $\underline{x} = \begin{bmatrix} x(1) \\ x(2) \end{bmatrix}$, $\underline{\mu} = \begin{bmatrix} \mu(1) \\ \mu(2) \end{bmatrix}$, $\underline{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$.

$$\langle x(i)|\xi\rangle = \mu(i) \tag{D.3}$$

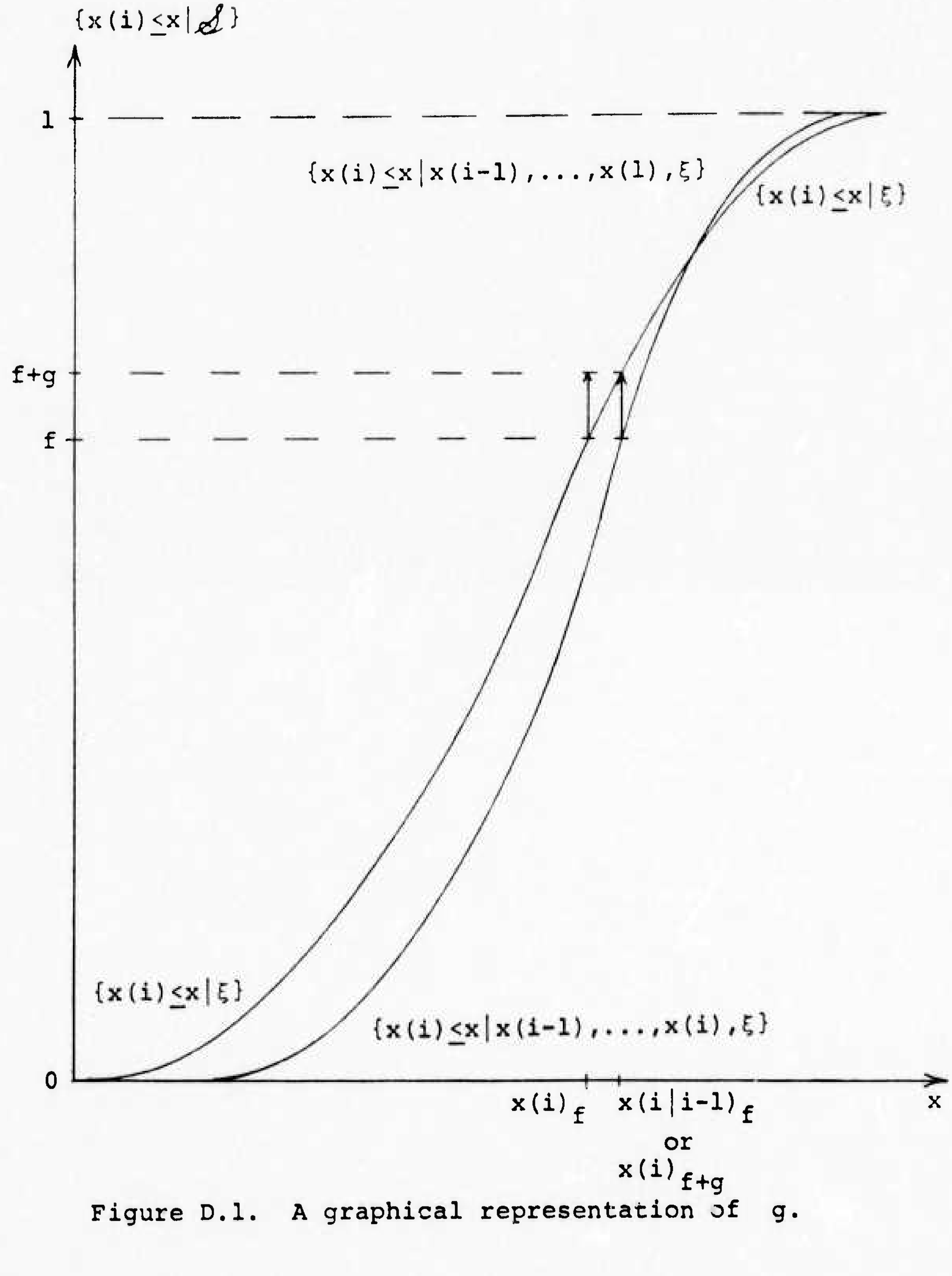


Figure D.1. A graphical representation of g.

$$x(i|\xi)_{0.5} = \mu(i) \tag{D.4}$$

$$\overset{v}{\langle} x(i)|\xi\rangle = \sigma_{ii} = \sigma(i)^2 \tag{D.5}$$

$$\sigma_{ij} = \rho_{ij}\sigma(i)\sigma(j) = \sigma_{ji}$$

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}} \tag{D.6}$$

$$\{x(2)|x(1),\xi\} = \left(\frac{1}{2\pi}\right)^{\frac{1}{2}} P_{22} \exp\left[-\frac{1}{2}P_{22}(x(2)-x(2|1)_{0.5})^2\right] \tag{D.7}$$

where $\underline{P} = \underline{\Sigma}^{-1} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$

$$\begin{aligned}
x(2|1)_{0.5} &= \langle x(2)|x(1),\xi\rangle \\
&= \mu(2) + \frac{\sigma_{21}}{\sigma_{11}}(x(1)-\mu(1)) \\
&= \mu(2) + \rho_{12}\sigma(2)\left[\frac{x(1)-\mu(1)}{\sigma(1)}\right] \\
&= \mu(2) + \rho_{12}\sigma(2)y(1)
\end{aligned}$$

$$\begin{aligned}
P_{22} &= \left(\sigma_{22} - \frac{\sigma_{12}^2}{\sigma_{11}}\right)^{-1} \\
&= \left[\sigma(2)^2(1-\rho_{12}^2)\right]^{-1}
\end{aligned}$$

$$\overset{v}{\langle} x(2)|x(1),\xi\rangle = \sigma(2)^2(1-\rho_{12}^2).$$

Finally for the normal distribution note that

$$x_f = x_{0.5} + y_f S_x. \qquad \text{(D.8)}$$

Therefore for the prior $\{x(2)|\xi\}$ and posterior $\{x(2)|x(1),\xi\}$ distributions on $x(2)$

$$\begin{aligned} g(f,y(1)) &= F_n\left[\frac{x(2|1)_f - x(2|\xi)_{0.5}}{S_{x(2|\xi)}}\;\middle|\; 0,1\right] - f \qquad \text{(D.9)} \\ &= F_n\left[\frac{x(2|1)_{0.5} + y_f S_{x(2|1)} - x(2|\xi)_{0.5}}{S_{x(2|\xi)}}\;\middle|\; 0,1\right] - f \\ &= F_n\left[\rho_{12} y(1) + y_f (1-\rho_{12})^{\frac{1}{2}} \;\middle|\; 0,1\right] - f\,. \end{aligned}$$

First note that $g(f,y(1))$ is restricted to the parallelogram shown in Figure D.2(a). If $x(1)$ and $x(2)$ are independent $(\rho_{12}=0)$, then g equals zero. If $x(1)$ and $x(2)$ are completely dependent $(\rho_{12} = 1 \text{ or } = -1)$, then g equals $f(1)-f$. This is shown in Figure D.2(b). Finally, note that when $f(1)$ equals zero $(y(1) = -\infty)$ or one $(y(1) = \infty)$, g equals zero. By examining Figure D.1, it is clear that these three relationships hold for any posterior and prior distributions.

In order to gain some insight into what g as a function of f for given values of $y(1)$ looks like, I have generated the following figures. Figure D.3 depicts g

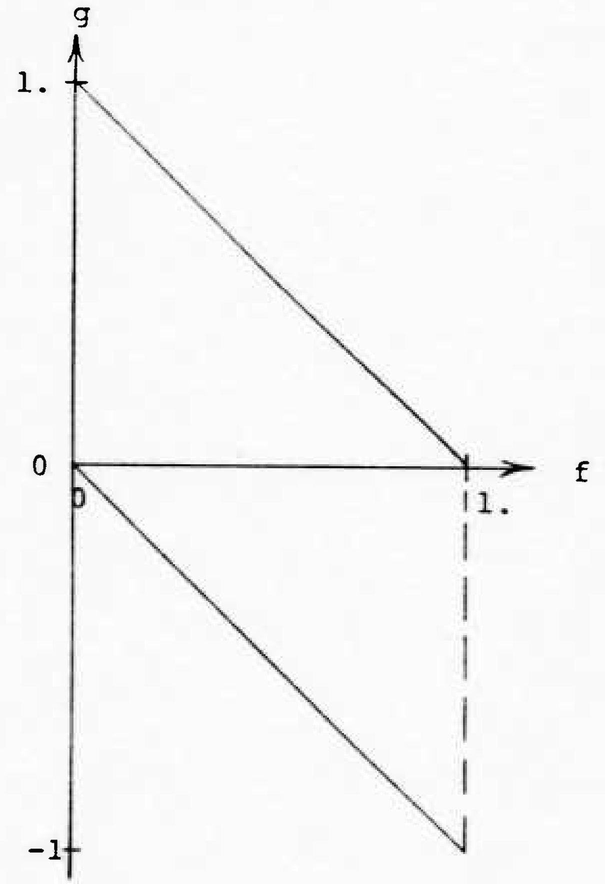


Figure D.2a. The bounds of g_f

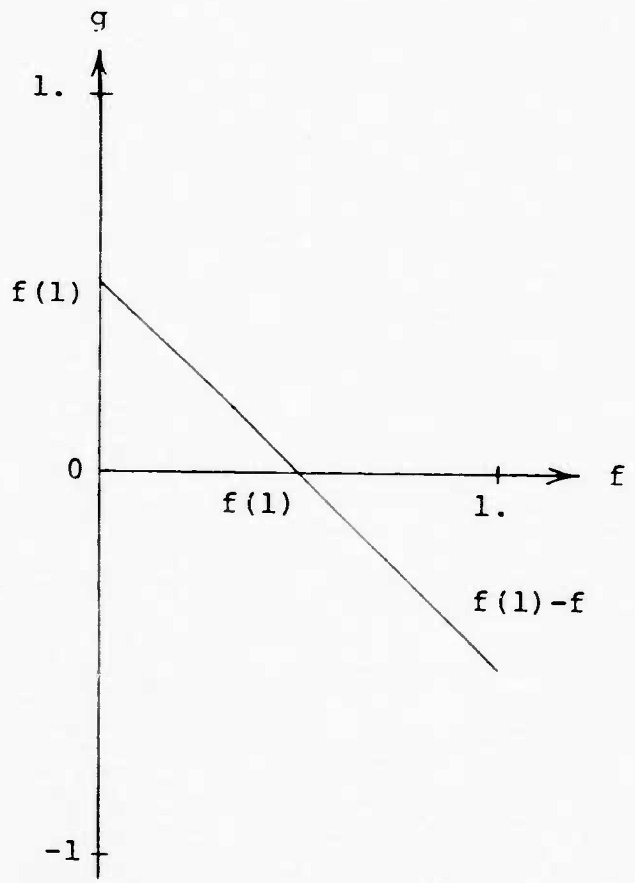


Figure D.2b. g_f when $\rho_{12} = \pm 1$

Figure D.2. Some properties of g_f

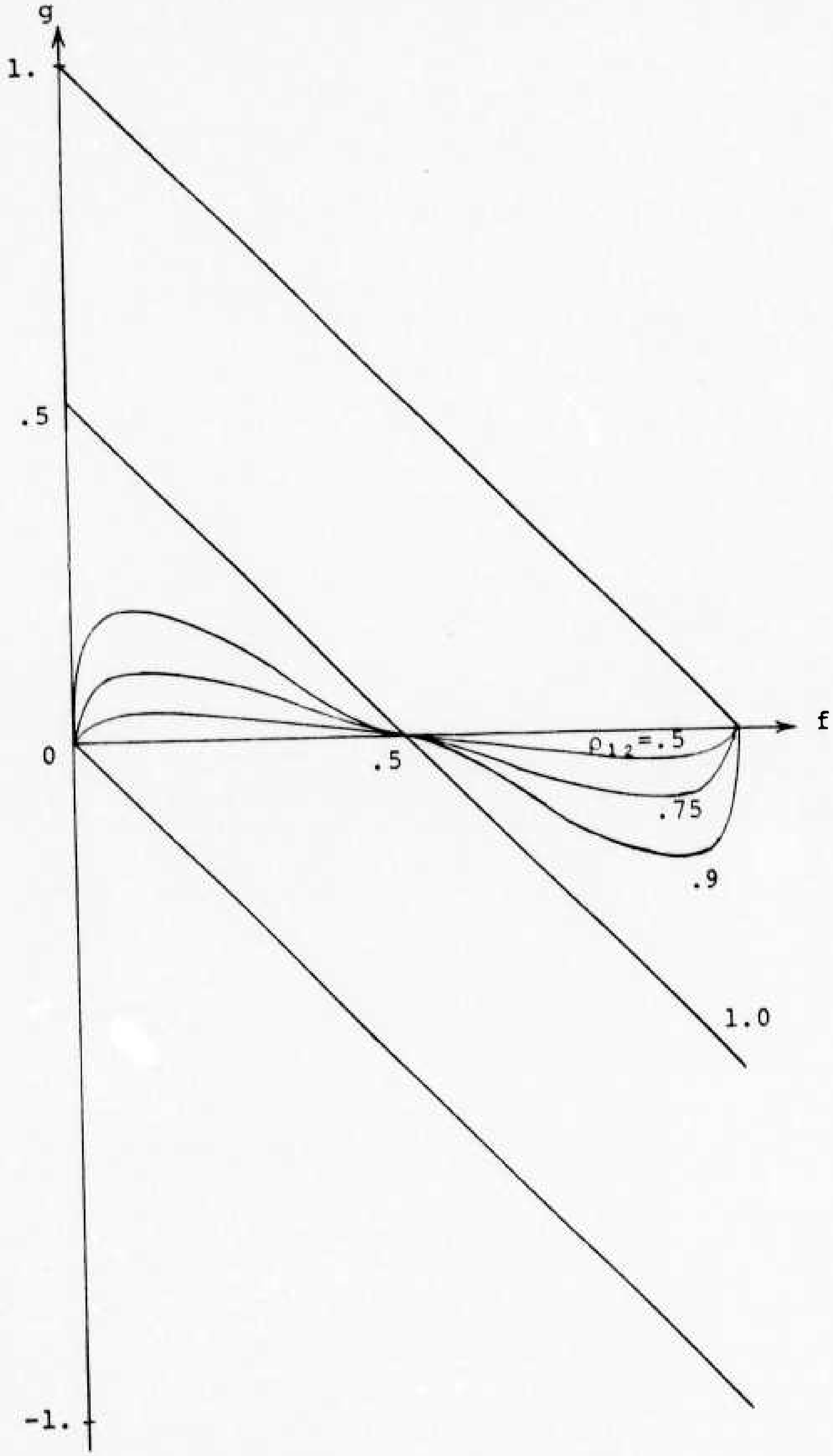


Figure D.3. A graph of g versus f for the bivariate normal, f(1) = .5

versus f for ρ_{12} equal to .5, .75, .9, and 1.0. It is very interesting to note that the curve for ρ_{12} equal to 0.9 is only about halfway between the curves for ρ_{12} equal to zero and one. Graphs of g versus f for $y(1)$ equal to .25 and ρ_{12} equal to .6, .8, and 1.0 are shown in Figure D.4. Examining this figure closely and comparing it to Figure D.3, one can see that as $f(1)$ moves away from 0.5, g gets closer to the limiting straight line. $f(1)-f$, for a given value of ρ_{12}. Examining equation (D.9) reveals that for $f(1)$ equal to zero or one g does equal $f(1)-f$, the limiting case. This is because y is unbounded.

Before leaving this example, we should make one more observation. If $x(1)$ and $x(2)$ were transformed to $z(1)$ and $z(2)$ as follows:

$$x(1) = \ln(z(1)) \tag{D.10}$$
$$x(2) = \ln(z(2)),$$

then $z(1)$ and $z(2)$ are described by a multivariate lognormal distribution, which is described in Chapter 2.3. However, $g(f,y(1))$ is equivalent to the g for the multivariate distribution, and so it is invariant to this change of variable.

This can be shown to hold for any posterior and prior distribution when the variable undergoes a monotonic

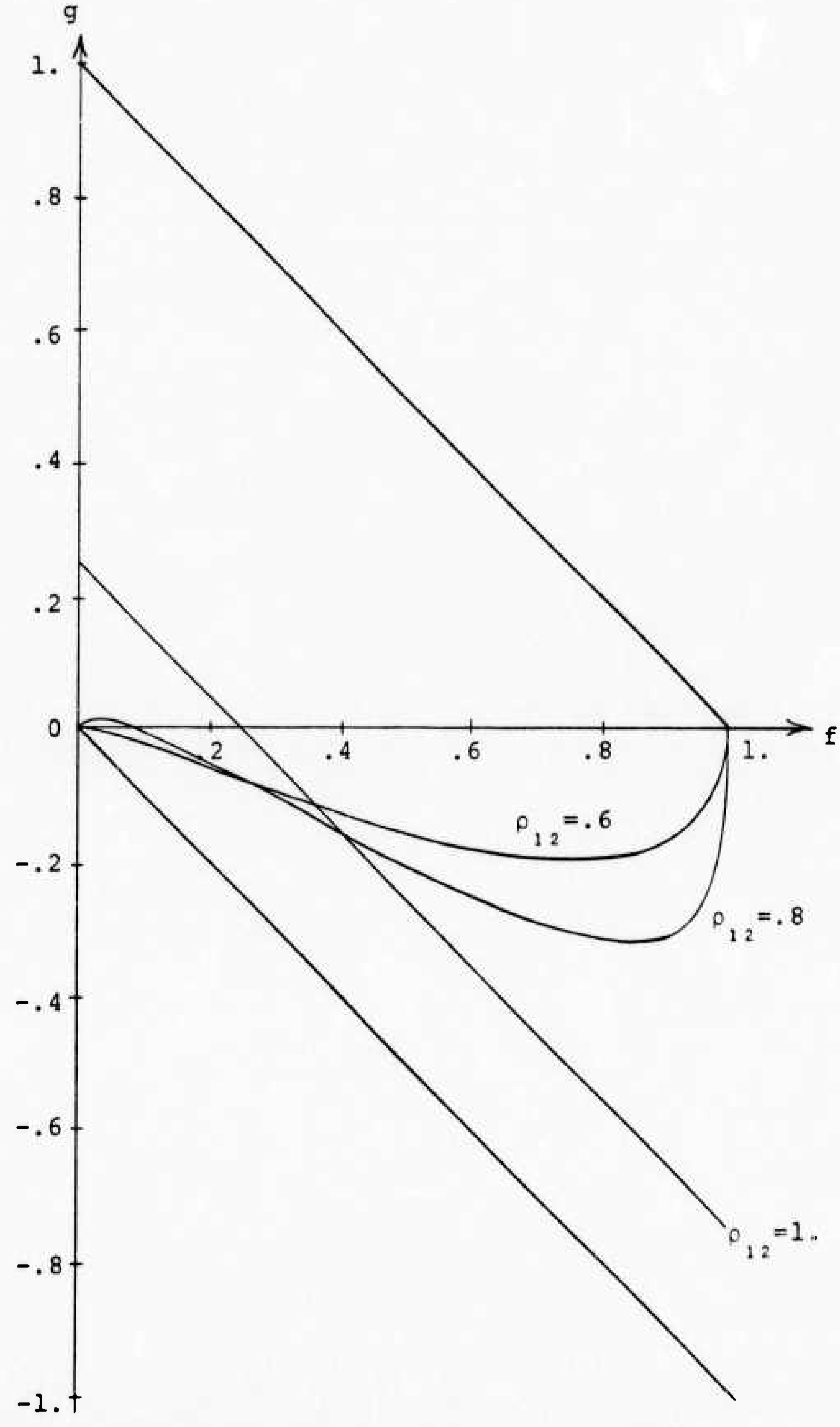


Figure D.4. A graph of g versus f for the bivariate normal, f(1) = .25

transformation using the graphical techniques for change of variable discussed in [11]. However, it does not hold for non-monotonic transformations.

Now let us examine g for the general method described in Chapter 3. We will do this for the case that all of the prior and posterior distributions are normal distributions. Let us assume that we have defined $g_{0.5}$ and $g^*_{0.9}$ to be functions of $f(1)$ for i equal to two. We have the following relationship from above since all of the distributions are normal:

$$g(f,f(1)) = F_n\left[\frac{x(2|1)_{0.5}+y_f S_{x(2|1)}-x(2|\xi)_{0.5}}{S_{x(2|1)}}\middle|0,1\right]-f. \tag{D.11}$$

Using equations 3.3 and 3.12 we get

$$g(f,f(1)) = F_n\left[F_n^{-1}(0.5+g_{0.5}|0,1)+\frac{y_f}{y_{0.9}}F_n^{-1}(0.9+g^*_{0.9}|0,1)\middle|0,1\right]-f. \tag{D.12}$$

This equation can be seen to satisfy all of the general properties mentioned above.

APPENDIX E

PROOF OF EQUATIONS 3.21 AND 3.23

The general form of equations 3.21 and 3.23 is

$$x(i|i-1)_f = F_{ln}^{-1}(f+g_f|\mu(i|\xi),\sigma(i|\xi)).$$

Solving for $(f+g_f)$ yields

$$f+g_f = F_{ln}(x(i|i-1)_f|\mu(i|\xi),\sigma(i|\xi))$$

$$= F_n\left[\frac{\ln(x(i|i-1)_f) - \mu(i|\xi)}{\sigma(i|\xi)}\middle|0,1\right]$$

by equation 3.19. Now solving for the argument of the normal cumulative and rearranging

$$\frac{\ln(x(i|i-1)_f) - \ln(x(i|\xi)_{0.5})}{\sigma(i|\xi)} = F_n^{-1}(f+g_f|0,1)$$

or $x(i|i-1)_f = x(i|\xi)_{0.5}(\exp(\sigma(i|\xi))\ F_n^{-1}(f+g_f|0,1))$.

REFERENCES

1. Aitchison, J. and J.A.C. Brown, The Lognormal Distribution, University of Cambridge Press, 1963.
2. Bather, J.A., "Invariant Conditional Distributions," Annals of Mathematics and Statistics, Vol. 36(3), 1965, pp. 829-846.
3. Bellman, R., Dynamic Programming, Princeton University Press, 1957.
4. Bhat, U. Narayan, Elements of Applied Stochastic Processes, J. Wiley & Sons, Inc., 1972.
5. Buede, Dennis, "The Use of a Linear Stochastic Model to Approximate the Joint Probability Distribution on an Evolving Dynamic Uncertain Process," unpublished paper, December, 1975.
6. Buede, Dennis, "A Sufficient Statistic Method for Approximating the Joint Probability Distribution on an Evolving Dynamic Uncertain Process," unpublished paper, February, 1976.
7. Chen, Chi-Tsong, Introduction to Linear System Theory, Holt, Rinehart & Winston, Inc., 1970.
8. Chu, John T., "On Bounds for the Normal Integral," Biometrika, Vol. 42(1955), pp. 263-265.
9. deFinetti, Bruno, Probability, Induction, and Statistics, J. Wiley & Sons, Inc., 1972.

10. DeGroot, Morris, H., Optimal Statistical Decisions, McGraw-Hill, 1969.

11. Gardner, Martin, "The Hierarchy of Infinities and the Problems it Spawns," in Infinity, ed. by William L. Schaaf, Board of Trustees of the Leland Stanford Junior University, 1969.

12. Georgescu-Roegen, Nicholas, The Entropy Law and the Economic Process, Harvard University Press, 1971.

13. Howard, Ronald A., "Class Notes for EES 221," unpublished.

14. Howard, Ronald A., "The Foundations of Decision Analysis," IEEE Transactions on Systems Science and Cybernetics, Vol. SSC-4, September, 1968, pp. 211-219.

15. Howard, Ronald A., "Proximal Decision Analysis," Management Science, Vol. 17(9), May, 1971, pp. 507-541.

16. Johnson, N.L. and S. Kotz, Distributions in Statistics: Continuous Multivariate Distributions, J. Wiley & Sons, Inc., 1972.

17. Jones, R.M. and K.S. Miller, "On the Multivariate Lognormal Distribution," The Journal of the Industrial Society, Vol. 16(2), pp. 63-76.

18. Kalman, R.E., "A New Approach to Linear Filtering and Prediction Problems," Journal of Basic Engineering, Vol. 82, 1960, pp. 35-45.

19. Karlin, Samuel and Howard M. Taylor, A First Course in Stochastic Processes, Academic Press, 1975.

20. Kyburg, Henry E., Jr., The Logical Foundations of Statistical Inference, D. Reidel Publishing Co., 1974.

21. Lempers, Frederick B., Posterior Probabilities of Alternative Linear Models, Rotterdam University Press, 1971.

22. Luenberger, David G., Optimization by Vector Space Methods, J. Wiley & Sons, Inc., 1969.

23. Meschkowski, Herbert, Evolution of Mathematical Thought, Holden-Day, Inc., 1965.

24. North, D. Warner, "The Abnormality of Normal Inferential Models for Processes that Evolve in Time," Proceedings of the Fifth International Conference on Operational Research, Venice 1969, Travistock Publications, London, 1970.

25. Raiffa, Howard and Robert Schlaifer, Applied Statistical Decision Theory, Harvard University Press, 1961.

26. Rao, Potluri and Roger LeRoy Miller, Applied Econometrics, Wadsworth Publishing Company, Inc., 1971.

27. Savage, Leonard J., The Foundations of Statistics, J. Wiley & Sons, Inc., 1954.

28. Spetzler, C.S. and C.S. Stael von Holstein, "Probability Encoding in Decision Analysis," in Readings in Decision Analysis, Decision Analysis Group, Stanford Research Institute, Menlo Park, California, 1974.

29. Tversky, Amos and Daniel Kahneman, "Judgement under Uncertainty: Heuristics and Biases," Science, Vol. 185, September, 1974, pp. 1124-1131.

30. Weber, Jean Draper, Historical Aspects of the Bayesian Controversy, University of Arizona Press, 1973.

31. Zellner, Arnold, An Introduction to Bayesian Inference in Econometrics, J. Wiley & Sons, Inc., 1971.

CONTRACT DISTRIBUTION LIST
(Uclassified Technical Reports)

Director Advanced Research Projects Agency Attention: Program Management Office 1400 Wilson Boulevard Arlington, Virginia 22209	2 copies
Office of Naval Research Attention: Code 455 800 North Quincy Street Arlington, Virginia 22217	3 copies
Office of Naval Research Attention: Code 102IP 800 North Quincy Street Arlington, Virginia 22217	6 copies
Office of Naval Research Branch Office, Boston 495 Summer Street Boston, MA 02210	1 copy
Defense Documentation Center Attention: DDC-TC Cameron Station Alexandria, Virginia 22314	12 copies
DCASMA Baltimore Office Attention: Mr. K. Gerasim 300 East Joppa Road Towson, Maryland 21204	1 copy
Director Naval Research Laboratory Attention: Code 2627 Washington, DC 20375	6 copies

SUPPLEMENTAL DISTRIBUTION LIST
(Unclassified Technical Reports)

Department of Defense

Assistant Director (Environmental and Life Sciences)
Office of the Deputy Director of Defense Research and Engineering (Research and Advanced Technology)
Attention: COL Henry L. Taylor
The Pentagon, Room 3D129
Washington, DC 20301

Director, Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209

Director, Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209

Defense Intelligence School
Attention: Professor Douglas E. Hunter
Washington, DC 20374

Department of the Navy

Office of the Chief of Naval Operations (OP-987)
Attention: Dr. Robert G. Smith
Washington, DC 20350

Office of Naval Research
Operations Research Programs (Code 431)
800 North Quincy Street
Arlington, VA 22217

Office of Naval Research
Operations Research Programs (Code 434)
800 North Quincy Street
Arlington, VA 22217

Office of Naval Research (Code 436)
Attention: Dr. Bruce MacDonald
800 North Quincy Street
Arlington, VA 22217

Office of Naval Research
Information Systems Program (Code 437)
800 North Quincy Street
Arlington, VA 22217

Director, ONR Branch Office
Attention: Dr. Charles Davis
536 South Clark Street
Chicago, Il 60605

Director, ONR Branch Office
Attention: Dr. J. Lester
495 Summer Street
Boston, MA 02210

Director, ONR Branch Office
Attention: Dr. E. Gloye
1030 East Green Street
Pasadena, CA 91106

Office of Naval Research
Scientific Liaison Group
Attention: Dr. M. Bertin
American Embassy - Room A-407
APO San Francisco 96503

Dean of Research Administration
Naval Postgraduate School
Attention: Patrick C. Parker
Monterey, CA 93940

Department of the Army

U.S. Army Research Institute
Organizations and Systems Research Laboratory
Attention: Dr. Edgar M. Johnson (PERI-OK)
1300 Wilson Boulevard
Arlington, VA 22209

Department of Engineering
United States Military Academy
Attention: COL A. F. Grum
West Point, NY 10996

Other Institutions

Director, Social Science Research Institute
University of Southern California
Attention: Dr. Ward Edwards
Los Angeles, CA 90007

Perceptronics, Incorporated
Attention: Dr. Amos Freedy
6271 Variel Avenue
Woodland Hills, CA 91364

Stanford University
Attention: Dr. R. A. Howard (10 copies)
Stanford, CA 94305

Decision Analysis Group
Stanford Research Institute
Attention: Dr. Miley W. Merkhofer
Menlo Park, CA 94205

Decision Research
1201 Oak Street
Eugene, OR 97401

Department of Psychology
Hebrew University
Attention: Dr. Amos Tversky
Jerusalem, Israel

Professor Howard Raiffa
Morgan 302
Harvard Business School
Harvard University
Cambridge, MA 02163

Professor Richard Meyer
Morgan 302
Harvard Business School
Harvard University
Cambridge, MA 02163

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: EES DA-77-5
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): THE ASSESSMENT OF A JOINT PROBABILITY DISTRIBUTION ON AN EVOLVING DYNAMIC UNCERTAIN PROCESS
5. TYPE OF REPORT & PERIOD COVERED: Technical
6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Dennis Melvin Buede
8. CONTRACT OR GRANT NUMBER(s): Contract N00014-76-C-0074; DDI Subcontract 75-030-0713; NSF Grant ENG 76-81056
9. PERFORMING ORGANIZATION NAME AND ADDRESS: The Board of Trustees of the Leland Stanford Junior University, c/o Office of Research Administrator, Encina Hall, Stanford, CA 94305
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Advanced Research Projects Agency, Human Resources Research Office, Arlington, VA 22209 (Subcontract from Decisions & Designs, Inc., McLean, VA 22101)
12. REPORT DATE: September 1977
13. NUMBER OF PAGES: 151
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Engineering Psychology Programs, Code 455, Office of Naval Research, 800 N. Quincy Street, Arlington, VA 22217
15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release. Distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

A dissertation submitted to the Department of Engineering-Economic Systems and the Committee on Graduate Studies of Stanford University in partial fullfillment of the Requirements for the Degree of Doctor of Philosophy, July 1977.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Decision Analysis
Probability Assessment
Dynamic Programming
Dynamic Random Variable
Joint Probability Distribution
Multivariate Named Distribution
Nonstationary Process

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An evolving dynamic uncertain process is one of several complicating factors in the analysis of decisions for which time is a critical element. A process is uncertain when its output cannot be specified deterministically; dynamic when it and its output change with the passage of time; and evolving when the output is revealed as time progresses. The decision-maker's uncertainty on the dynamic random variable that characterizes such a process must be adequately represented in any decision analysis. The direct assessment of this uncertainty is preferable to further modeling when the expert's

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

knowledge of the variable is roughly equivalent to his knowledge of its components. This dissertaion addresses the problem of performing the assessment of a joint probability distribution on a dynamic random variable in a practical manner.

First, multivariate named distributions are examined and found to be too rigid in their structure to be good approximations of the decision-maker's joint probability distribution, in general.

The major segment of this dissertation presents a method of assessing a joint probability distribution on a dynamic random variable. This method is a general model of the decision-maker's use of new information to update his probabilistic beliefs. A mathematical framework for specifying a posterior distribution in terms of the prior distribution and the revealed information is introduced here. This framework provides the analyst with a great deal of flexibility for approximating any joint distribution the decision-maker might have.